Considerations for the Montana Aligned to Standards Through-year (MAST) Program in Science

July, 2025

The Montana Science Assessment Task Force
With Support From:
The National Center for the Improvement of Educational Assessment

Prepared for the Montana Office of Public Instruction

Subtitle

The National Center for the Improvement of Educational Assessment, Inc. (the Center for Assessment) is a New Hampshire-based not-for-profit (501(c)(3)) corporation. Founded in September 1998, the Center's mission is to improve student learning by partnering with educational leaders to advance effective practices and policies in support of high-quality assessment and accountability systems. The Center for Assessment does this by providing services directly to states, school districts, and partner organizations to support state and district assessment and accountability systems.

SUGGESTED CITATION:

D'Brot, J., Marion, S., & Dadey, N. (2025, August). Assessment Design and Implementation Considerations for the Montana Alignment to Standards Through-Year Program in Science. Dover, NH: The National Center for the Improvement of Educational Assessment. Available online: [url]

Table of Contents

Executive Summary

The Montana Office of Public Instruction (OPI) has launched a comprehensive redesign of the state's science assessment system, building on recent innovations in English language arts and mathematics. This initiative seeks to develop a next-generation, through-year science assessment system that is instructionally useful, fair, and responsive to local contexts.

To guide this work, OPI convened a statewide Science Assessment Task Force in spring 2025, bringing together educators, district leaders, content specialists, and policy experts. Through a series of virtual and in-person meetings, the Task Force developed and refined a theory of action (TOA) to anchor the design and implementation of the new system. This collaborative effort emphasized the importance of deep, real-world scientific reasoning, a coherent vision for student learning across grade spans, and support for meaningful instructional practice in all Montana classrooms.

The resulting MAST (Montana Aligned to Standards Through-Year) Science Theory of Action outlines a framework for an assessment system designed not only to measure learning, but also to support it. It focuses on aligning instruction and assessment with Montana's Science Standards, providing multiple points of feedback throughout the year, and equipping educators and families with timely, actionable information. The TOA reflects Montana's commitment to ensuring that every student, regardless of location or background, has access to high-quality science instruction and the opportunity to demonstrate what they know and can do.

The Montana Theory of Action (TOA)

The Science Assessment Task Force brought deep knowledge of Montana's educational landscape and a strong commitment to ensuring equitable learning opportunities for all students. Guided by the Center for Assessment, the Task Force conducted a series of structured meetings to define a shared vision for science learning in Montana and identify the most important outcomes that students should achieve. Their work was grounded in Montana's science content standards.

The Task Force co-developed a TOA that serves as both a strategic vision and a practical roadmap for a next-generation science assessment system. The TOA outlines how the assessment will operate within a through-year model, supporting instruction at multiple points during the school year rather than relying on a single end-of-year test. It clarifies the roles of the state, districts, and schools in supporting implementation, and emphasizes the importance of accessible data, local curricular flexibility, and teacher agency. It identifies the core outcomes Montana wants for its students, including scientific reasoning, critical thinking, and the ability to apply knowledge to real-world and culturally relevant problems. It also links these outcomes to concrete assessment and instructional practices. Importantly, the TOA presents an idealized set of conditions, actions, and responsibilities that can help articulate what needs to be in place for the system to work at its best.

Key Components of the System

The TOA outlines how Montana's science assessment system is intended to function, rather than how it currently operates. For this vision to become a reality, every level of the education system must work in collaboration, with each partner understanding and fulfilling their distinct role in supporting high-quality science learning.

Below are the core components of the MAST Science system:

- Clear performance expectations tied to grade-level outcomes in grades 5, 8, and high school
- A shared vision of what science learning should look like across Montana's classrooms
- A distributed, through-year design that supports instructional pacing and reduces test-day pressure
- Deliberate attention to supporting small, rural, and under-resourced districts through shared tools, guidance, and implementation supports
- Reporting systems that return meaningful, user-friendly information to educators, students, and families

In addition to the system's design, this report offers detailed recommendations to ensure successful implementation. These include establishing a design and implementation advisory group, investing in educator capacity-building, creating clear project plans and timelines, and prioritizing the design of reporting systems that reinforce learning. The report also outlines considerations for piloting the new system, including how feedback from districts and educators will be used to inform future iterations.

Together, the theory of action and accompanying recommendations provide a coherent, forward-looking blueprint for improving science education in Montana, anchored in the belief that thoughtful assessment design can deepen learning, support teaching, and promote equity statewide.

Introduction and Background

The Montana Office of Public Instruction (OPI) received a U.S. Department of Education grant to design a science "through-year" assessment system. A through-year assessment program is a statewide assessment program that is (1) administered in multiple distinct sessions during a school year, and (2) intended to support the production and use of a summative determination, and one or more additional goals. Following its previous work in ELA and mathematics, OPI engaged the National Center for the Improvement of Educational Assessment (Center for Assessment) to support the development of a framework for the MAST Science assessment system. As part of this initiative, the Center facilitated a task force of Montana educators and leaders to co-develop a theory of action that would anchor design decisions and clarify the system's goals, structures, and intended outcomes.

This report outlines the process undertaken by the task force and presents the proposed theory of action. This theory of action is intended to guide the design, implementation, and ongoing evaluation of the Montana Aligned to Standards Through-Year (MAST) Science assessment system. It offers both a strategic framework and practical guidance for realizing a next-generation, instructionally useful, and locally relevant assessment system.

From the outset, the Task Force recognized the value of grounding its work in a clear, shared vision. In its initial meetings, the group began defining the purpose of the science assessment, naming key instructional outcomes, and exploring the feasibility of a through-year model. These early sessions helped clarify the roles of the state, districts, and schools in supporting student learning. Collaborative activities and draft visuals helped the group co-develop a preliminary Theory of Action (TOA), which served as a conceptual anchor for design discussions. This report presents a summary of the deliberations and recommendations of the Task Force.

The Task Force

OPI leadership invited a diverse and knowledgeable set of representatives to participate in the task force, as seen in the table below.

Table 1. Task Force Membership

Name	Affiliation	Role
Norah Barney	Anaconda School District	Curriculum Director
Jeannie Chipps	Montana State University- Science Math Resource Center, Department of Education	Higher Education
Carissa Jenkins	Fairfield Public Schools	Middle School Teacher
Erin Lynch	Havre Public Schools and MSU-Northern	Elementary School Teacher

Patrick McClellan	Bozeman School District 7	Middle School Principal
Andrea Meiers	Lockwood Schools	Curriculum Director
Tobin Novasio	Hardin Schools	District Superintendent
Chris Pavlovich	Ripple: The Center for Education and Ecosystem Studies	Higher Education
Karen Pollari	Sidney Public Schools	High School Teacher
Linda Rost	Baker Public Schools	High School Teacher
Sarah Shannon	Ashland Public School	District Superintendent
Billi Taylor	School Services of Montana	Curriculum Director
Tim Tharp	Board of Public Education, Chair	State School Board
Monica Tomayer	Conrad School District	Middle School Teacher
Rebecca Turk	Montana State University	Higher Education

In addition to Juan D'Brot, Nathan Dadey, and Scott Marion from the Center for Assessment, the Task Force was supported by Julie Murgel, Krystal Smith, Erin deGraw, Michelle McCarthy, and Cedar Rose from OPI.

Task Force Process and Meetings

The Task Force meetings started in early March and concluded with a final meeting on June 20, 2025. These included four virtual meetings and one in-person convening. Each session was designed to surface challenges with the current system, establish shared design principles, and refine the system's instructional goals and implementation strategies.

During the in-person convening in May 2025, the task force reviewed and revised an early version of the theory of action. Discussions focused on the intended teaching and learning outcomes, instructional and logistical constraints, and the reciprocal roles of the state, district, and school. The group also responded to draft visuals and design ideas proposed by members. In addition to producing a theory of action, the task force engaged in design activities to help bring the theory of action to life. These design activities were intended to elicit the task force members' assessment design preferences, such as the ways in which the standards would be bundled into testlets, the number of testlets administered each year, and the grades in which the testlets would be administered.

Following the in-person meeting, the task force reconvened virtually to review a refined narrative, revised visuals, and use cases for assessment implementation. Feedback focused on ensuring accessibility for small districts, the role of curriculum consortia, and supports for family and community engagement. These refinements informed the third and final version of the MAST Science Theory of Action, which is documented in this report.

MAST Background

Following a year of piloting in 2022-2023 and then field testing in 2023-2024, the Montana Aligned to Standards Through-Year (MAST) assessment program in ELA and Mathematics was first administered in the 2024-2025 school year. Summative scores will be used in the federal accountability system for identifying Targeted Support and Improvement (TSI), Additional Targeted Support and Improvement (ATSI), and Comprehensive Support and Improvement (CSI) only for small schools after standard setting in the summer of 2025. CSI for all other schools will occur in the Fall of 2026.

To extend the MAST model into science, MT OPI was awarded a Competitive Grant for State Assessments from the U.S. Department of Education in late 2024. *MAST Science* is intended to support Montana in improving teaching and learning in elementary and middle school science and provide educators with a more comprehensive assessment system in these grades. Specifically, the OPI envisions *MAST Science* serving student and educator needs by (1) improving science proficiency across all student groups by leveraging timely, actionable assessment data to promote standards-aligned instruction and (2) reflecting Montana's cultural heritage, constitutional requirement and duly enacted policy commitment to Indian Education for All (IEFA), the assessments will incorporate science practices and perspectives of Montana's American Indian communities.

MAST Science's planned assessment framework will focus on Indigenous cultural understandings in the Montana Science Content Standards (Montana Board of Public Education, 2016). For example, the standards include topics such as ethnobotany (cultural uses of plants), cultural fire management, and traditional technologies. Including IEFA items aligned to , standards also deepens the coherence between the MAST Science assessment and instruction.

The work of the Task Force produced more than a series of meetings; it created a shared vision and a foundation for action. These insights from the Task Force serve as a bridge between assessment design principles and the broader context of Montana schooling. Understanding the state's demographics, geography, and instructional realities is critical for a system intended to be both feasible and impactful. The following section situates the MAST Science design within the unique educational landscape of Montana.

Context of Montana Schooling

Montana's 826 public schools serve approximately 150,000 students, supported by roughly 11,000 teachers. Nearly all school districts (about 96%) are classified as small and rural. At last count, there are 428 elementary, 225 middle, and 173 high schools in the state. Given that

Montana's population of 1,137,233[1] is distributed over almost 146,000 square miles, over half of Montana's schools enroll fewer than 100 students, including 65 one-room schoolhouses[2]. This geographic reality shapes nearly every aspect of schooling, from staffing and curriculum offerings to assessment administration and family engagement. Only 51 schools serve more than 500 students. These conditions create significant variability in educator capacity, course offerings, and resource access, which the MAST Science system must respect and support.

Demographic Characteristics

While the majority of Montana students identify as White (77%), the state has a significant Native American population (10%), representing 12 tribal nations across seven reservations. Montana's Indian Education for All is designed to ensure all schools have the knowledge, tools, and resources necessary to honor the IEFA requirement and integrate it into their teaching materials and methods.

The demographic characteristics of other students include:

- 6% Hispanic students
- 5% multiracial
- Fewer than 1% in each of the African American, Asian, and Pacific Islander categories
- 14% students with disabilities (special education)
- 2.4% English learners
- Slightly fewer than 20% of students are eligible for free or reduced-price meals

These population characteristics highlight the importance of equitable access to both instruction and assessment, as well as the need for plain-language, family-friendly reporting that supports engagement across diverse communities. The following section explores how these contextual factors translate into differences in instructional time, curriculum organization, and classroom experiences; differences that the MAST Science assessment system must be designed to accommodate.

Variability of Science Instruction

Science instruction in Montana varies considerably across grade levels, schools, and districts, reflecting the diverse local contexts in which teachers operate. Several factors contribute to this variability:

[1] Population estimate as of July 1, 2024 from the U.S. Census Bureau: https://www.census.gov/quickfacts/fact/table/MT/PST045224
[2] From https://www.montana.edu/crre/overview.html

- **Elementary science is often limited and integrated**. In many K–5 settings, teachers are generalists who juggle multiple subjects and have limited dedicated time for science instruction. In some cases, science is delivered only a few times per week, or embedded into English language arts or social studies units.
- **Middle school instruction varies between domain-based and integrated approaches**. Some districts adopt a domain-based sequence—life, earth, and physical science—while others use integrated courses aligned to NGSS-like progressions.
- **High school offerings differ widely by school size and resources**. Larger schools may offer a full progression of biology, chemistry, physics, and earth science, while very small or rural schools may rotate course offerings annually or combine grade levels to deliver required content.
- **Local curriculum decisions drive differences in scope and sequence.** Some districts leverage shared curriculum consortia to maintain coherence and access to instructional materials.

These variations directly impact how students experience science and when they encounter key concepts, which in turn influence how and when they can demonstrate mastery.

For the MAST Science assessment system, this variability reinforces the Task Force's recommendations for:

1. Flexible testlet sequencing that accommodates different instructional approaches and pacing.
2. Through-year administration that supports timely feedback without overburdening single test windows.
3. Designing tasks that are instructionally useful across both integrated and domain-based curricula.

By acknowledging these differences in science instruction, the assessment system can remain fair, locally relevant, and feasible for all districts.

Variability of Resources and Capacity

Resource availability and educator capacity vary widely across Montana schools, and these differences have direct implications for assessment implementation and data use.

- **Educator roles in small schools are often wide-ranging**. In many rural and one-room schoolhouses, teachers take on multiple subjects and grade levels, and administrators or instructional coaches often support several functions. This breadth of responsibility can make it more challenging to dedicate time for activities like data analysis, curriculum alignment, and professional learning.

- **Instructional resources reflect local context.** Laboratory equipment, science kits, and digital tools are often more readily available in larger districts, while smaller or remote schools may adapt local resources creatively to support hands-on science experiences. Technology infrastructure, including broadband access, can vary, which is an important consideration for both digital instruction and online assessments.
- **Professional learning opportunities can be less accessible in remote areas**. Teachers in small or geographically isolated schools may have fewer chances for in-person collaboration or state-level training. This highlights the value of regional curriculum consortia and virtual professional learning networks, which can connect educators across distances and support shared capacity-building.

The MAST Science system must be designed with these capacity realities in mind. Assessment tools, reporting, and professional learning supports should be:

1. Accessible to educators with multiple responsibilities, minimizing additional administrative burden.
2. Supported by state- or regionally shared resources, including sample tasks, pacing guides, and curated instructional supports.
3. Accompanied by virtual and asynchronous professional learning, allowing educators in remote areas to access training and collaborate across districts.

Addressing variability in resources and capacity ensures that the benefits of MAST Science are equitably distributed and that even the smallest, most rural schools can engage in data-informed instruction and continuous improvement cycles envisioned in the Task Force's Theory of Action.

These differences in instruction, resources, and educator capacity illustrate why the MAST Science system must be flexible and responsive to local contexts. Rather than expecting a single approach to meet the needs of all schools, the Task Force emphasized the importance of designing a system that builds on Montana's strengths (small, community-oriented schools, local curriculum decisions, and place-based learning), while still providing consistent, meaningful information for all students. To anchor this system in a shared vision, the Task Force first identified the student learning outcomes that should guide science teaching and assessment statewide. The next section presents the Task Force's vision for what Montana students should know and be able to do, forming the foundation for both curriculum alignment and assessment design.

Prioritized Outcomes

One of the early activities undertaken by the Task Force was the identification of key student learning outcomes that should guide science instruction and assessment across the state. Members sought to define not just what students should know, but how they should think, reason, and engage with the world scientifically. The resulting prioritized outcomes emphasize scientific literacy, reasoning, and application across students' daily lives and future pathways. These outcomes are meant to reflect a vision of science learning that builds over time—progressing from foundational understanding in early grades to sophisticated reasoning and application by high school. The outcomes also reflect Montana's values around place-based learning, fairness, and preparing students to be informed citizens capable of making evidence-based decisions in diverse contexts.

The prioritized outcomes, followed by specific expectations for students at the end of grades 5, 8, and high school are presented below.

Reason Scientifically About Their Lives and the World Around Them. Successful students can reason scientifically and act as critical thinkers and problem solvers. Students can locate and evaluate evidence, including considering and evaluating conflicting viewpoints. After students have received high-quality opportunities to learn science, they will be scientifically literate, will be able to successfully reason about scientific phenomena, and will be able to wisely evaluate and consume a range of information. These students can then use their scientific understanding to make informed decisions about their lives, e.g., in terms of their health, professions, and the world.

Show Increasingly Sophisticated Scientific Reasoning Across Grades. Expectations should increase as development increases, rooted in the standards/NGSS progressions. For example, K-5 students learn science at the macro level, before shifting to the micro level in middle school, and finally, they learn to work across macro and micro levels in high school. This also means that background knowledge from previous grade levels should be used to work toward more advanced understandings.

Draw on and apply core understandings. While the standards are non-negotiable, some ideas are more important depending on the specific areas of study. The three-dimensional science was designed to help students develop deep understandings of three intertwined domains: disciplinary core ideas, science and engineering practices, and cross-cutting concepts. Students should be able to draw on all three domains to reason scientifically about real-world phenomena.

In addition to these high-level outcomes across grades, the Task Force identified the following outcomes corresponding to expectations for students at the end of grades 5, 8, and high school:

Table 2. Student Expectations by the end of Target Grades

Grade	Students will know and be able to do the following…
By the end	• Apply science and engineering practices to simple applications involving 2-3

of 5th grade..	variables. ● Develop foundational scientific literacy, including the ability to ask questions, construct explanations, and use basic models. ● Understand macro-level science concepts, such as ecosystems, forces, and earth systems, before transitioning to micro-level concepts in later grades. ● Demonstrate early skills in evaluating evidence and making claims based on multiple sources of data. ● Develop an initial understanding of human impacts on the environment, especially place-based phenomena and begin to interpret data related to such phenomena.
By the end of 8th grade…	● Apply more sophisticated use of the science and engineering practices, incorporating multiple variables and systems while still working with familiar scientific concepts. ● Strengthen their ability to analyze and interpret data, recognizing patterns and relationships. ● Demonstrate increasing critical thinking skills, using reasoning and argumentation to support claims. ● Develop a more progressively complex understanding of scientific knowledge and skills across disciplines. ● Develop early competence in engineering and design thinking, such as developing and using models and simulations to solve problems, particularly place-based phenomena. ● Transition to micro-level scientific reasoning, making connections between atomic/molecular-level processes and observable phenomena.
By the end of high school…	● Demonstrate mastery of open-ended scientific applications, independently using science and engineering practices to design experiments and analyze complex data. ● Read, interpret, and critique scientific journals and studies to make informed decisions about health, careers, and civic responsibilities. ● Apply claim-evidence-reasoning (CER) frameworks to assess conflicting scientific viewpoints. ● Exhibit strong data literacy, recognizing bias, assumptions, and the reliability of multiple sources. ● Understand and evaluate the ethical implications of scientific advancements. ● Navigate seamlessly between macro- and micro-level scientific reasoning. ● Engage in problem-solving with real-world applications, demonstrating scientific curiosity and the ability to identify and investigate new place-based problems and more general phenomena.

These grade-level expectations serve a dual purpose. First, they provide a developmental roadmap to guide science instruction from elementary through high school, reinforcing Montana's commitment to coherence, relevance, and fairness across learning experiences. Second, they inform key aspects of the assessment system's design, including the selection of tasks, types of reasoning emphasized, and the nature of feedback provided to students and educators.

In future stages of assessment development, these expectations may also be used to shape performance level descriptors (PLDs), guide the development of scoring rubrics, and anchor student exemplars. Task Force members envisioned these expectations as supporting clear communication with educators, students, and families about what "good science learning" looks like at each stage. However, before they are used to shape PLDs, the task force recommends vetting these with a broad-based group of science educators to make sure these grade-level expectations are complete and appropriate.

As the system moves from design to pilot implementation, these learning expectations will serve as a foundation for curriculum alignment, professional learning supports, and student-centered reporting. Grounding the assessment design in real-world applications and scientific thinking ensures that Montana's students are not only assessed on what they know, but on how they use that knowledge to navigate and shape their world.

Constraints and Requirements

Assessment design is an exercise in optimization under constraints. Therefore, Task Force members spent a fair amount of time outlining the likely constraints and requirements under which MAST science must operate. The requirements specify what the assessment must do. In general, the requirements outline how the scores will be used and for what purposes. Constraints, on the other hand, outline the challenges or considerations under which the assessment must fulfill these uses and purposes. For example, if a requirement called for the assessment to reliably measure how students conduct real-world investigations and reason with evidence, but the constraints indicated that the assessment cannot take more than two hours to administer, designers would have to wrestle with this conflict. Describing the constraints and requirements can surface these tensions so the designers can try to create the most optimal solution possible.

Requirements

Task force members generated many potential requirements. As can be seen, the Task Force members indicated that the test must be required to meet key technical, accessibility, and participation requirements. Additionally, Task Force members recommended that the science assessment must support instructional uses and measure student longitudinal growth.

1. Under current law, state assessments must produce annual summative determinations (e.g., proficiency) for each student.
    a. These must be comparable across students, subgroups, schools, and time.
2. The statewide achievement tests must be aligned, as fully as possible, to the state's learning standards.
3. All students must participate in the state assessment, and the state must ensure that the assessment is accessible and fair to all students.

4. The assessment must measure the full range of the score continuum to allow students to show growth.
5. The assessment must also address the Indian Education for All (IEFA) state requirement.
6. The test results must be accurate and actionable.
7. All test items must be of high quality, developed using principal assessment design approaches.
8. All students must be assessed at least once during elementary and middle school[3].
9. The assessment must be designed to allow students to show growth.
10. The assessment must allow students from a variety of multiple standards-aligned curricula to fairly demonstrate their knowledge of science concepts and skills.

Constraints

The Task Force used a similar process to identify the potential constraints that might limit the potential designs. These range from things like testing time, resource issues, and curricular flexibility. The list of identified constraints is presented below.

1. MAST science cannot require noticeably more testing time than the current Montana science assessment.
2. The implementation itself must not require substantially more time for teachers (e.g., multiple hours of training in how to use the test administration system).
3. High levels of student transience and absences will make it challenging to ensure that all students are participating in all or enough testlets to produce a valid overall score.
4. Limited staff availability in some districts may require creative scheduling for make-up testing. This should not require extraordinary efforts from school staff members.
5. Local technological capacity varies considerably and must be taken into consideration when designing the assessment.
6. Must "fit in" with the testing time for current MAST and NAEP (8th grade) testing.
7. Student motivation is a potential constraint if the testlets are administered long after or before instruction in that subdomain has occurred.
8. Local control and the local flexibility required to implement this assessment in Montana are constraints on the potential comparability of scores across districts.
9. Local education agency capacity, both in terms of people and finances, is a constraint that must be considered in the design and implementation of MAST science.
10. OPI must have the capacity to implement the system and support local district personnel well.
11. Montana's large geographic area and considerable number of small schools are constraints in terms of supporting schools all across the state.

3The purview of the Task Force did not include high school. State are required to assess science at least once during the high school grade span.

Current MAST testing time

The Task Force was especially concerned with the testing time of the MAST math and ELA as a specific constraint. The testing time for the math and ELA tests is shown below.

	Math	**ELA**
Number of testlets	12	6 + 2 performance tasks
Test Windows	3	3
Testlets/Window	3-5	2
Minutes/Testlets	30	30
Total Testing Time	**6 hours + T13**	**4.5 hours**

Given the amount of testing time allocated for mathematics and ELA, the Task Force was wary of recommending more science testing than necessary. Nevertheless, the Task Force recommended designing a test long enough to meet the key requirements and design priorities, but no longer.

Assessment Design Considerations

A theory of action is a critical tool for guiding the design and evaluation of a program or policy. The theory of action developed for the MAST language arts and mathematics through-year assessment program has steered MT OPI and its testing company partner as they have developed and implemented this assessment system. The theory of action presented here for the new science assessment was designed to serve similar purposes.

Theories of action can be challenging for those not initially involved in creating the theory of actions to understand because graphics can only convey so much information. To address some of these challenges, the facilitators engaged the task force in high-level design activities to clarify their views on the potential science assessment designs.

The task force acknowledges that the proposed designs may not be the eventual design of the next Montana science assessment. Rather, the task force contends that these potential designs help bring the theory of action to life and can help support the more detailed design activities yet to come.

Overview of a Design Framework

The task force spent considerable time deliberating potential assessment designs. While there were differences of opinion, there was essentially universal agreement about the following five key principles for the design of Montana's "through-year" science assessment:

1. The assessment should employ a testlet design to assess the required content for each grade span: K-5, 6-8, and high school. Each testlet should be constructed to measure student learning of specific standards or appropriate groupings of standards.
2. The testlets should be designed to accommodate both integrated and domain-based curriculum approaches.
3. The sequencing of administering individual testlets should largely be under the control of district administrators.
4. Federal law requires the administration and reporting of science assessment scores at least once at the elementary, middle, and high school levels. However, the task force recommends administering testlets in each of the following grades: 3, 4, and 5 for elementary school, and in grades 6, 7, and 8 for middle school. This recommendation is tied to the task force's interest in ensuring that science is taught and assessed at all grades in order to help support a coherent program of science instruction in each school.
5. The score of record should be based primarily on the scores from the testlets administered in the terminal grade (i.e., 5, 8, and 12) in the particular grade span.

We expand on each of these five points below.

Standards and Testlets

When Montana adopted its state science standards, it adopted the "performance expectations" from the Next Generation Science Standards as its content standards. Performance expectations combine the three strands of science education—the disciplinary core ideas, the science and engineering practices, and the cross-cutting concepts—into statements of what students should know and be able to do by the end of each grade span. Many states have used a similar approach when adopting standards in hopes of making the potential combinations of disciplinary core ideas (the content), science and engineering practices, and cross-cutting concepts more manageable. Nonetheless, there are still a considerable number of standards for each grade span.

There are 75 performance expectations in elementary school and 57 in middle school. Task force members acknowledged that it would be challenging to meaningfully assess all these standards at each grade span. Grouping the 75 standards into a manageable number of testlets, such as 12, would result in more than six standards included on each testlet. With nine testlets, which could work better for reasons discussed below, it would result in more than eight standards on each testlet. One of the primary objectives of the new assessment system is to provide teachers and students with instructionally useful information. To do this, each standard should be assessed with multiple items (e.g., six or more) to shed light on the degree to which students have learned the standards and to gain insights into where they might need additional support. The problem of too many standards is a little less challenging at middle school, but this would still require assessing between six and ten standards on each testlet.

The task force struggled with how to address this challenge. One recommendation is for OPI to convene a group of science educators and other stakeholders to prioritize which standards should be assessed and how often each standard should be assessed. Some states have done this by designating certain standards as “assessment standards” and others as “instruction standards.” OPI could provide classroom assessment tasks and other assessments for locals to use in assessing students against the instruction standards.

Alternatively, OPI can rotate standards, such that not all standards are assessed in each testing cycle. Determining how standards are rotated in and out involves many comparability, cost, and communication challenges. Rotation can accomplish the goal of assessing all standards, just not annually. Again, this often happens with conventional testing programs.

These are just a couple of suggestions. The task force recommends that OPI collaborate with its constituents to prioritize the standards and assessment framework so that the assessment system is both manageable and effective.

Testlets and Curriculum Grouping

The task force was clear that the science assessment must produce valid and fair scores, regardless of whether students learned science through an integrated or a domain-based approach to the curriculum. This will require producing essentially two different assessments, which will have significant cost and comparability implications. Therefore, if OPI follows these recommendations, it will likely limit other potential design options. One approach could involve creating mini-testlets based on individual standards or very small groups of standards and then configuring testlets differently depending on the curricular orientation.

Critically, the scores from both sets of assessments must be comparable at the student and school levels. This will require designing comparability studies, likely by embedding items from one assessment format into the other and vice versa. This will be one of the more challenging aspects of the science assessment design.

Flexible Testlet Administration

The task force recommended allowing district leaders to determine the sequence of testlets administered in their schools. There will be no perfect sequence given the wide variety of science curriculum scope and sequence in Montana. Task force members suggested that there could be even more variability in the integrated approach compared with the domain-based approach to science. Nevertheless, the task force recommended constructing testlets and suggesting sequencing to try to match the most commonly used curriculum sequencing. The Task Force made this recommendation to support local control of science curriculum and instruction, while also recognizing that some degree of state guidance could support districts that lack deep science expertise.

Administration Across Grade Levels

The task force expressed a strong preference for administering the testlets across grades within a grade span for 3-5 and 6-8. The technical experts identified several challenges in attempting to test across three grades, including computing an overall score when students move in and out of the system. The task force members, based on their experience, recommended that the benefits of spreading the testing across grades outweighed the challenges raised by the technical consultants. Some of these benefits included ensuring that science was taught at every grade level, and reducing the amount of testing needed at any one grade.

The task force noted that the grade level standards in elementary grades (3, 4, & 5 explicitly) provide a direct connection to standards to assess, whereas middle school standards are grade-band standards. Thus, distributing the tests across three grades may be easier at the elementary school. This means the testlets must be interchangeable across grades for comparability reasons and to manage costs. Task force members recognized the differences in general academic preparation and maturity across these grade levels. Therefore, they recommended that, to the extent possible, the reading level associated with the testlets should be targeted to the lowest grade in the span or even lower if possible. While this may be possible, task force members recommended collecting validity evidence to evaluate the degree to which comparability is achieved.

The task force recognized that the science assessment must operate in the context of the English language arts and mathematics assessments. ELA and math testlets are administered in three windows across the school year. Therefore, it makes sense for science to follow the same pattern. The desire to adhere to the ELA and math schedule, combined with the recommendation to administer assessments across each grade span, led the task force to suggest administering one science testlet in each window for a total of nine testlets across the grade span.

Scores of record

Computing science scores of record for Montana students would be challenging under any through-year design. These challenges are exacerbated when testlets are administered across multiple grade levels. Even though mobility rates may not be as high in Montana as in other states, they can add up faster than one might think. Furthermore, even if students stay in Montana but move from one district to another, they will likely have completed a different set of testlets in their original district than those in their new district.

Additionally, combining scores from testlets administered across three grade levels raises concerns about the comparability and validity of the results. Nobody doubts that sixth and eighth-grade students, for example, are at very different levels of maturity and academic preparation. It is risky to assume that the scores from a testlet administered to sixth and eighth graders mean the same thing and requires considerable evidence to substantiate claims of score interchangeability. However, testlets in sixth and seventh grade are not used for accountability. They would be leveraged as instructional tools and feedback mechanisms with state reporting taking place at the eighth grade level only.

Therefore, the task force urged OPI to consider alternative scoring models that rely on scores from the previous grade. The three testlets administered in fifth or eighth grade would include at least 36 total items (perhaps more), which could be enough to produce a reliable score. The task force also discussed including at least one performance task that would require students to demonstrate their ability to synthesize knowledge and skills across multiple standards. If such a task were included in the assessment program, it would enhance the validity and credibility of the overall score.

Prioritizing the scores from the terminal grade when producing a total score does not preclude using results from the earlier grades as part of the score, but would require a different scoring model than the typical "gradebook" approach. Rather than recommending a specific approach for computing total scores, the task force strongly urges OPI and the assessment vendor to conduct research into the scoring models described here, as well as other potential models.

A Vision for Montana students

From the outset, the MAST Science Task Force recognized the value of grounding its work in a clear, shared vision. In the initial virtual meetings, the group began identifying the foundational components necessary for a well-aligned science assessment system. These early sessions focused on defining the purposes of the assessment, establishing desired outcomes, and clarifying the roles of various system levels in supporting student learning. Principally, the following vision emerged as a result of early meetings:

> *Montana's scientifically capable and competent high school graduates require an assessment system that aligns with developmental progressions, critical thinking, and real-world applications. This vision must address the challenges of technology access, curriculum variability, teacher capacity, and assessment logistics.*

Using draft visuals and structured collaborative activities, Task Force members co-developed an initial version of the Theory of Action (TOA), which served as a conceptual foundation for design discussions. This early version outlined key inputs, processes, and outcomes, and helped clarify the respective roles of the state, districts, and schools. While the initial framing followed a largely linear structure to support shared understanding, it laid the groundwork for iterative refinement. Feedback from both virtual and in-person meetings led to subsequent revisions and a more robust TOA described in later sections of this report.

Revisiting the Theory of Action

After the in-person Task Force convening, which included extended working sessions focused on practical design elements, the draft TOA was revisited for review and refinement. Members engaged with revised visuals and suggested that a deeper narrative be developed to support a reader's interaction with the TOA. This subsequent version reflected key themes raised during the in-person meeting, including the need for a more cyclical representation of the system, the role of place-based learning, the importance of clearly identifying teaching outcomes, and the constraints imposed by limited instructional time in some schools, especially in grades K-5.

The second version of the TOA was presented during a virtual feedback session where Task Force members were asked to evaluate the narrative framing, visual clarity, terminology, and level-specific role descriptions. The Task Force provided robust feedback, affirming the revised TOA while suggesting important clarifications, such as expanding the role of curriculum consortia in rural areas, clarifying the distinction between technical and instructional support, and improving the accessibility of the visualizations. These inputs were synthesized and incorporated into the next iteration of the TOA. This is presented in the next section.

The MAST Science Theory of Action

The MAST Science Theory of Action (TOA) offers a comprehensive roadmap for how Montana's science assessment system is intended to support equitable, instructionally useful, and locally relevant learning across all levels of the system—state, district, school, and student. It is grounded in Montana's science standards and shaped by Task Force feedback, which emphasized clarity, system coherence, and feasibility of implementation in both small and large district contexts.

Overall Purpose of the Theory of Action

The Task Force reiterated the need to state the purpose and vision of the MAST Science regularly to schools, districts, and the community. At a high level, the purpose of MAST Science is to improve student learning in science and deepen understanding of scientific concepts through a coherent and supportive assessment system. Practically, the TOA is intended to highlight how the State of Montana, its districts, and schools should design supports to engage educators and administrators in working toward the vision presented in the previous sections.

Multiple Views of the Theory of Action

This Theory of Action is presented in multiple formats to serve different purposes:

- The High-Level TOA presents the vision of the system in a simplified form that communicates key goals and processes to the public and educators.
- A middle layer, the Overview of the Detailed TOA, helps orient readers to the structure and progression of ideas.
- The Detailed TOA expands on this vision to illustrate the interconnected, reciprocal roles of the state, district, and school—capturing the operational details that guide implementation, evaluation, and refinement.

While the high-level TOA was initially linear for communication clarity, feedback from the Task Force highlighted the cyclical nature of the work, including continuous engagement with data, reflection, and instructional adjustment are expected. The TOA is not a compliance mechanism. It is a tool to support ongoing improvement and shared understanding. It presents an idealized set of conditions, actions, and responsibilities that can help articulate what needs to be in place for the system to work at its best.

Several important themes from Task Force input have been incorporated into this narrative:

- The theory of action is grounded in the Montana Science Standards.
- Place-based learning is a goal for instruction but is not directly assessed; this distinction is important and should be supported with examples and aligned instructional resources from the state.
- There is a considerable challenge of limited instructional time across K-12 in science, with particular limitations in elementary settings and small districts, especially where a single teacher may assume multiple roles.
- Schools and districts can benefit from professional learning (both technical and instructional), delineating between "how to" administer the assessment and "how to" use the resulting data for instruction.
- The TOA should emphasize the cyclical nature of assessment and instructional improvement efforts, despite the visual limitations of a series of linear steps. This is emphasized by the through-year testing design, which offers multiple data points to guide instructional adjustments, interventions, and supports.
- Teaching outcomes (like dedicated science instructional time and data-informed practice) are critical to influencing student outcomes.

In the remainder of this section, we describe the role of the state, district, and school, define several terms that may be helpful to the reader, provide examples of roles and supports at each level, and offer several implementation considerations. After these sections, we then present the visual displays of an overall, high-level theory of action, as well as a more detailed theory of action.

The State, District, and School Roles

This section introduces "level descriptions." These are detailed role-specific narratives that help clarify what each level (state, district, school) is expected to do within the TOA, how each level supports the others, and what implementation challenges and practical examples might arise. By highlighting key responsibilities, implementation challenges, and real-world examples at each level, this section aims to foster a shared understanding of how coordinated action across the system leads to meaningful, sustained improvement. It is important to note that rather than working in silos, these descriptions illustrate how the three levels can align their efforts, support one another, and collectively make the assessment system happen.

- **State Level**. The state provides a framework for assessment, ensures alignment with state standards, and offers training and resources to districts, including guidance and curated examples for local curriculum alignment, district consortia coordination, and support for implementing the assessments with fidelity. Additionally, the state monitors and compares trends across districts to guide policy decisions and equal access initiatives.
- **District Level**. District leaders collaborate with school and state leaders to align curriculum and pacing guides with Montana standards, organize professional development, coordinate data use in schools, and facilitate community outreach about uniform reporting practices. Districts also bridge the gap between state guidance and school implementation, ensuring that teachers have clear expectations and support.
- **School Level**. Schools use MAST results to inform teaching adjustments and interventions, identify student needs, and engage in regular cycles of reflection and improvement supported by administrators and professional learning communities (PLCs). They also ensure that teaching time and data-informed instruction are balanced against the realities of testing time and instructional load, with administrators playing a key role in creating supportive conditions for teachers to implement data-informed practices.

Given these descriptions of the state, district, and school roles, we highlight how these responsibilities are interconnected and mutually supportive. The examples of professional development, resource sharing, and data use at each level show how leadership and instructional supports cascade from the state level down to classrooms. These interconnected roles reflect the collaborative, cyclical nature of the TOA. To help readers navigate these detailed frameworks and examples, we now shift to clarifying the language and technical terms used throughout this document.

Clarifying Key Language and Concepts

To ensure shared understanding across the state, this section defines specialized terms used throughout the Theory of Action that may not be immediately clear to all readers. It ensures that technical jargon is minimized and the language is as accessible as possible, so everyone can understand how the pieces of the TOA fit together.

Table 3. Key Language

Term	Description
Balanced Assessment System	Balanced assessment systems provide feedback to students and information for teachers to support rich instructional and learning opportunities. Assessments outside of the classroom, at the district and state levels, provide aggregate data to policymakers and education leaders, allowing for the monitoring of educational opportunities and support for high-quality instruction indirectly through the provision of appropriate curricular resources and professional development opportunities. The multiple assessments in the system are coherently connected via a shared vision of student learning (Marion, Pellegrino, & Berman, 2024).
Coherence	Ensuring that curriculum, instruction, and assessments align with shared goals and standards (NRC, 2006).
Comprehensiveness	Assessment provides a variety of evidence to inform educational decision-making (NRC, 2001).
Continuity	The ability of assessments to provide information that allows monitoring and evaluation over time (NRC, 2001).
Continuous Improvement Cycle	The process of regularly reviewing data and practices to make ongoing improvements
Curriculum Consortia	Groups of districts collaborating on curriculum alignment and resource sharing
Data-Informed Practice	Teaching decisions based on data.
Differentiation	Tailoring instruction to meet individual student needs.
Efficiency	This refers to getting the most out of assessment resources and eliminating redundant, unused, and unnecessary assessments (Marion, et al., 2019)
Equity	Equity can be defined in many ways. This can range from equity in outcomes, distribution of resources, access to experiences, expectations in growth, or resources and opportunities (Levinson, Geron & Brighouse, 2022).
Equity Initiatives	Efforts to ensure fair access and outcomes for all students.
Fairness	Fairness in assessment means that all test takers are treated equitably, and that test scores reflect only the intended constructs—not irrelevant

	factors like race, gender, disability, or socioeconomic status (AERA, APA, NCME, 2014).
Full Academic Year (FAY)	The idea that students must be in a school for a certain time to count in accountability.
Formative Assessment	Ongoing assessment practices that are used to adjust teaching and learning, not for grading.
Interim Assessment	Periodic assessments during the year to measure progress toward standards.
Performance Expectations	The specific skills and knowledge students are expected to demonstrate.
Place-Based Learning	Learning that draws on local contexts and resources.
Pupil-Instruction-Related (PIR)	The term for a day of teacher activities devoted to improving the quality of Instruction (see OPI, 2024).
Professional Learning Communities	Teacher teams working together to plan and improve instruction based on a careful examination of student assessment results and related evidence of student learning.
Summative Assessment	An assessment that evaluates student learning at the end of an instructional period.
Testlet	A single instance of an assessment, often as part of multiple testlets administered during a test window for a given content area.
Through-Year Testing	Multiple smaller assessments throughout the year rather than one large end-of-year test. They typically support accountability systems and at least one other aim (Dadey, Evans, & Lorie, 2023)
Utility	The degree to which assessments provide the information necessary to support their uses and purposes (Marion et al, 2019).

Examples of Roles and Supports at Each Level

This section introduces examples drawn from the task force that illustrate the roles and supports at each level of the system, making it easier for readers to understand how these responsibilities work in practice and why they matter.

State Level. At the state level, the Montana Office of Public Instruction (OPI) can work to support schools and districts and consider how to best address challenges across regions. The OPI currently hosts summer workshops and can build on the work by, potentially, providing online resources, such as videos of model lessons, and providing support to regional service centers. For example, the OPI might host a regional meeting where district leaders practice analyzing student work samples using the new science standards, and then share those insights with teachers in their districts. The state can also organize working groups to gather teacher feedback on assessment blueprints, adjusting them based on real classroom needs.

The following represent sample action steps that the state could conduct to support districts:

- Include sample assessment tasks or instructional supports that reflect Montana's diverse regions and natural resources.
- Host summer workshops and regional meetings.
- Provide videos of model lessons, sample tasks, and data interpretation guides applicable to different state and local contexts.
- Organize working groups to gather educator feedback on assessment blueprints and implementation plans.
- Develop and share implementation guides and communication toolkits tailored to a range of district needs, ensuring they are relevant to diverse district and school contexts.

District Level. At the district level, district leaders can coordinate with instructional coaches to create and adapt pacing guides, identify gaps in resources, and plan monthly check-ins with school principals. For example, a district might hold monthly curriculum review meetings, at least during the first year, to discuss science pacing guides and share best practices for integrating new tasks. They also might organize a back-to-school science assessment bootcamp for teachers, where district leaders could walk through sample tasks, discuss where students typically struggle, and co-develop strategies to support those areas.

District data coaches could also create easy-to-read dashboards for principals to see how their schools are performing on through-year assessments. Furthermore, districts could take advantage of curriculum consortia to supplement limited personnel or resources. For example, staff in small or rural districts may cover multiple roles or be limited to just one or two science teachers. Here, curriculum consortia may fulfill some of these responsibilities. These consortia often bridge gaps in capacity and serve as a proxy for district-level coordination.
The following represent sample action steps that districts might consider in a variety of contexts:

- Work with instructional leaders to create and adjust curriculum guides and pacing documents.
- Conduct professional learning opportunities to share and explore new resources and instructional practices.
- Facilitate ongoing discussions with school administrators and instructional teams to monitor progress and identify improvement areas.

- Offer sample instructional materials and collaborate on resource development to fill gaps, ensuring relevance for diverse district contexts.
- Facilitate family and community outreach about uniform reporting practices, including creating plain-language summaries of student results or hosting school-based family nights to review science performance data

School Level. At the school level, teachers may want to gather in professional learning groups to look at assessment results, share what works or doesn't in their lessons, highlight students' misconceptions, and co-design future lessons to address these challenges. These learning groups could include teacher-led horizontal or vertical teams, curriculum committees, data review committees, assessment use committees, or professional learning communities, or networked improvement communities. A fifth-grade teacher might use a mini-assessment from the state's science bank to check if students understand weather patterns and adjust their unit plan based on that data. School leaders, meanwhile, could work to ensure teachers have protected time to meet and share these insights and provide substitute coverage when needed for additional collaboration time.

The following represent sample action steps that school leaders may wish to incorporate as part of their work in using MAST Science:

- Gather teachers in small professional learning groups, PLCs, or other vertical and horizontal teams to review interim or periodic assessment data to identify student learning needs and adjust instruction.
- Use short, informal checks for understanding (like mini-assessments) to identify learning gaps and modify unit plans as needed.
- Organize teacher planning days or professional learning sessions, providing flexibility and coverage as needed, recognizing that some schools may have limited resources or staff.
- Facilitate targeted support and ensure dedicated time for collaborative planning, making sure this is feasible in smaller and larger school contexts alike.
- Communicate assessment results and instructional goals clearly to families through newsletters, open houses, or parent-teacher conferences.

This cross-level approach can help ensure schools are actively engaged in improving science learning in Montana.

Implementation Considerations

This section provides an overview of the implementation challenges associated with the TOA and the Task Force's feedback. These challenges highlight practical obstacles that schools and districts face, such as differences in capacity and funding across regions, difficulties in translating state-level guidance into daily practice, and balancing instructional demands with data-informed decision making. Table 4 (below) provides examples and mitigation strategies that offer practical insights and solutions to support successful implementation and continuous improvement.

Table 4. Implementation Considerations and Mitigation Strategies

Challenge	Example	Potential Mitigation Strategy
Limited PD time	Small schools may have insufficient dedicated days for professional development (PD).	Evaluate funding use to ensure the minimum number of hours dedicated to PD. Create virtual PD options and flexible after-school sessions.
Substitute coverage gaps	Rural districts struggle to find substitutes for teacher data days.	Share resources regionally or partner with nearby districts for coverage.
Data trust and use	Teachers need additional support to interpret and act on new data.	Provide data coaching and clear data-use protocols.
Balancing testing and instruction	Teachers feel overwhelmed by testing demands.	Build testing schedules collaboratively and communicate the purpose of interim assessments.
Fair access to opportunities and coherence	Variability in resources across schools.	Provide shared state-level resources and examples that work in diverse contexts.
Families and guardians may misunderstand data	Score descriptions may range from detailed reports to scores without explanation.	Develop state-provided or district-designed templates and communication toolkits for explaining scores and instructional implications.
Gaps in digital literacy or access	Some families may lack access to or understanding of digital score reports.	Provide printed summaries or hold in-person sessions at schools to walk through reports and data dashboards.

Lack of clarity on how families can support instruction	Families want to help, but are unsure what they can do with the results.	Include actionable next steps for families in score reports (e.g., sample questions to ask, home activities).

The implementation concerns associated with MAST Science include district capacity (especially in small or rural districts), limitations on professional development time, balancing instructional and testing demands, and building trust in the data to drive meaningful instructional change. For example, districts may struggle to provide substitute coverage for teacher collaboration days, and smaller schools may have limited staff to lead data analysis sessions. While there may be multiple ways to address these concerns, districts can consider collaborating with one another to share resources and staff, and state or regional agencies can provide additional coaching and support to build data fluency and reduce pressure on smaller districts.

Ultimately, many of the challenges will need to be addressed through a vertically integrated approach across levels of the system. Additionally, partnerships across districts and schools, supported by the state, may be necessary.

Visualizing the MAST Science Theory of Action

The high-level TOA represents the broader framework and major outcomes of the system, while the Detailed TOA offers detailed connections between state, district, and school roles and activities. Together, they illustrate how Montana's science assessment system is designed to guide continuous improvement and data-informed practice across all levels. They are intended to help readers understand both the big picture and the operational details of the MAST Science TOA.

High-Level Theory of Action

The high-level Theory of Action outlines the major goals and guiding principles of Montana's science assessment system. It shows how the system is designed to build trust in data use, promote data-informed decision-making, and foster continuous improvement cycles across all levels. This visualization details how the TOA connects with Montana's place-based learning goals, access considerations, and accountability structures, offering a broad vision that can be adapted to local contexts.

It describes how the state provides a vision and overarching support, while districts customize implementation based on local needs and schools take action in classrooms. While the high-level TOA is presented as three linear sections, the work should be conceptualized as cyclical. It emphasizes the iterative nature of data review and instructional changes support the growth of all learners. The high-level TOA provides a broad framework for how Montana's science assessment system works. It includes the key components, goals, and guiding principles that set the direction for implementation at the state, district, and school levels. This overview explains how the system aims to foster equitable learning opportunities, data-informed decision-making, and continuous improvement cycles. Key inputs include the state's role in establishing a clear vision and flexible design framework, targeted performance expectations, high-quality reporting, and supporting district implementation through training, curated examples, and data interpretation.

Detailed Theory of Action

The Detailed Theory of Action builds upon the high-level framework, specifying how each level contributes to the shared goals. It identifies the explicit roles of the state, district, and school in supporting teachers and students, sharing resources, and using data to inform instruction.

Throughout the report, we provided several examples, such as how districts might work with instructional leaders to develop pacing guides or how teachers might engage in PLCs to analyze data and adjust instruction. We also raise potential barriers, such as limited capacity in small schools. The detailed TOA expands on the high-level framework with detailed descriptions of the roles and responsibilities at the state, district, and school levels. For example, it highlights the importance of coordinating professional development to support teachers in interpreting data or sharing resources to address capacity limitations. It also acknowledges challenges such as balancing testing and instruction and fostering trust in data use, while suggesting that collaborative structures and shared resources can help mitigate these challenges.

Figure 1. Simplified Theory of Action for Montana's Science Assessment System.

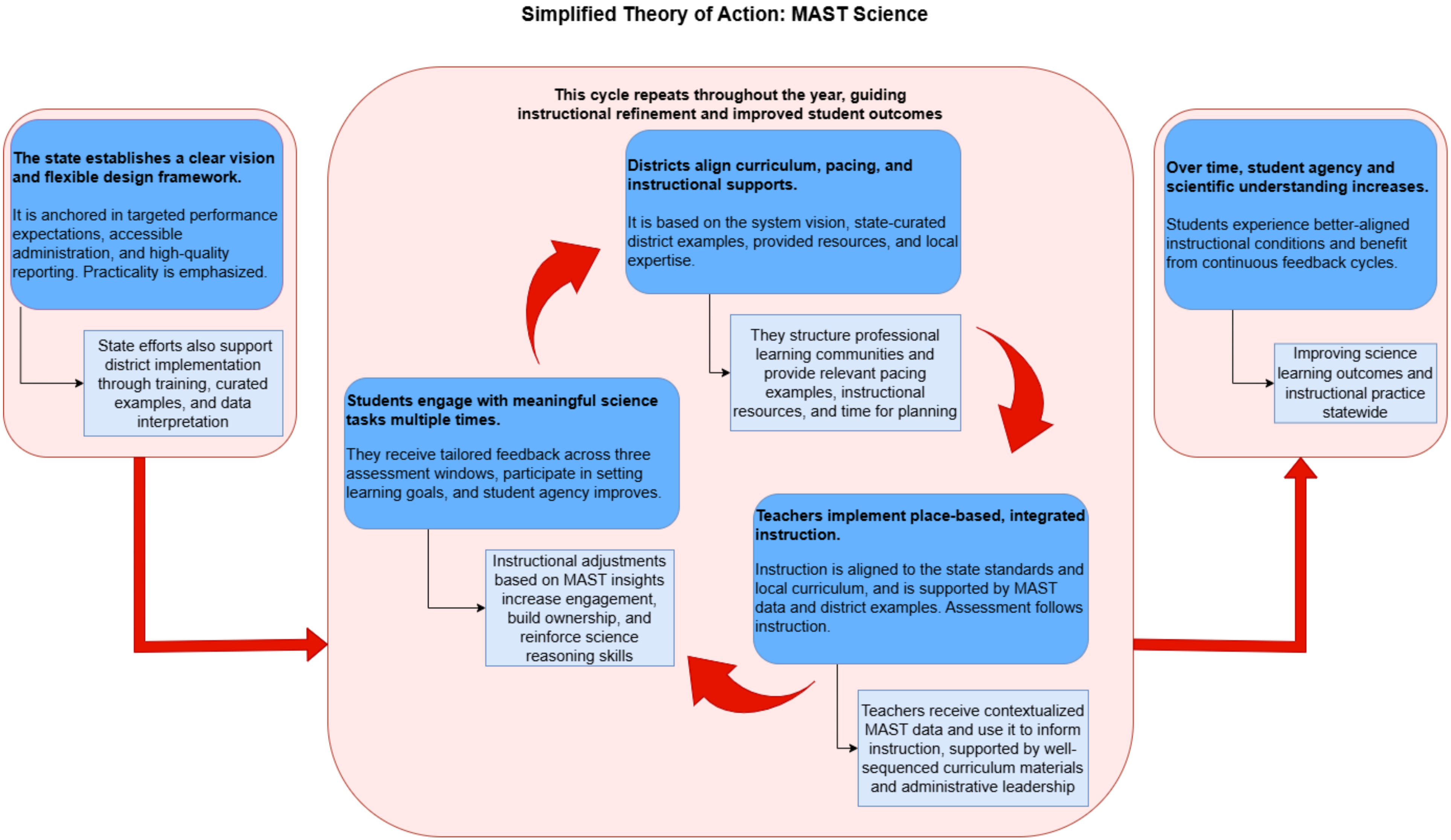


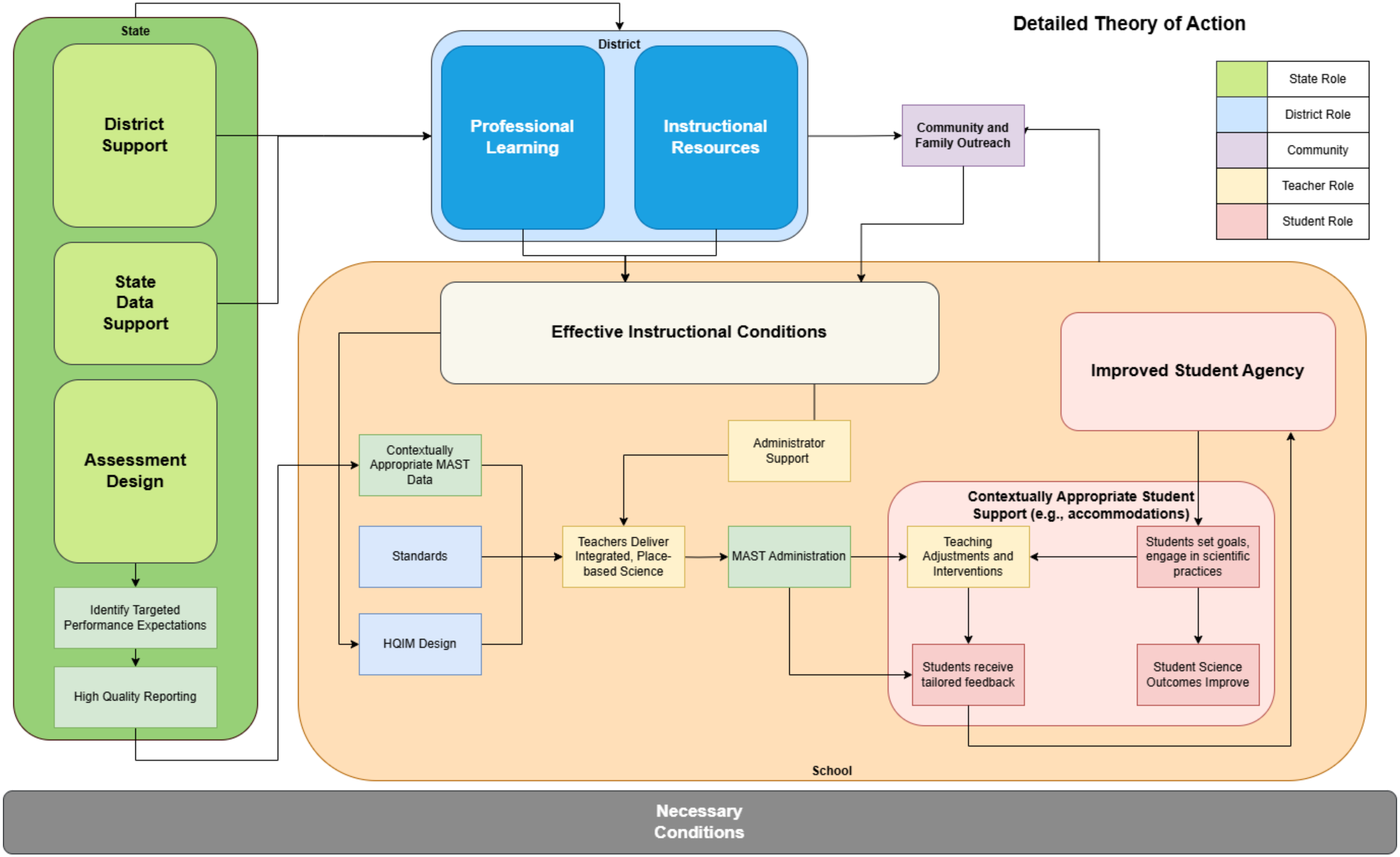
Figure 1 above represents the Simplified Theory of Action for Montana's Science Assessment System (MAST Science). This high-level visualization illustrates the state establishing a clear vision and provides a flexible design framework. This includes setting targeted performance expectations, offering practical supports such as curated examples and training, and promoting accessible administration and actionable reporting. These state-level efforts create the foundation upon which districts build.

Districts then align curriculum, pacing, and instructional supports with the state's vision and resources, using professional learning communities and local resources to meet school and district needs. This alignment enables teachers to implement place-based, integrated instruction that reflects both state standards and local contexts. Teachers are supported by sequenced curriculum materials, MAST data, and instructional leadership.

Within this instruction and assessment system, students engage with meaningful science tasks multiple times throughout the year. Feedback from these experiences fosters student agency, deepens understanding, and allows for timely instructional adjustments based on real-time data insights. As the system repeats this cycle of engagement and refinement, student agency and scientific understanding increase over time. To increase student agency, the assessment and learning system must support high-quality feedback in ways that allow students to reflect on their own learning and growth.

The end goal is to improve science learning outcomes and instructional practice statewide through a coherent, data-informed, and locally responsive system.

Figure 2. Detailed Theory of Action for Montana's Science Assessment System.



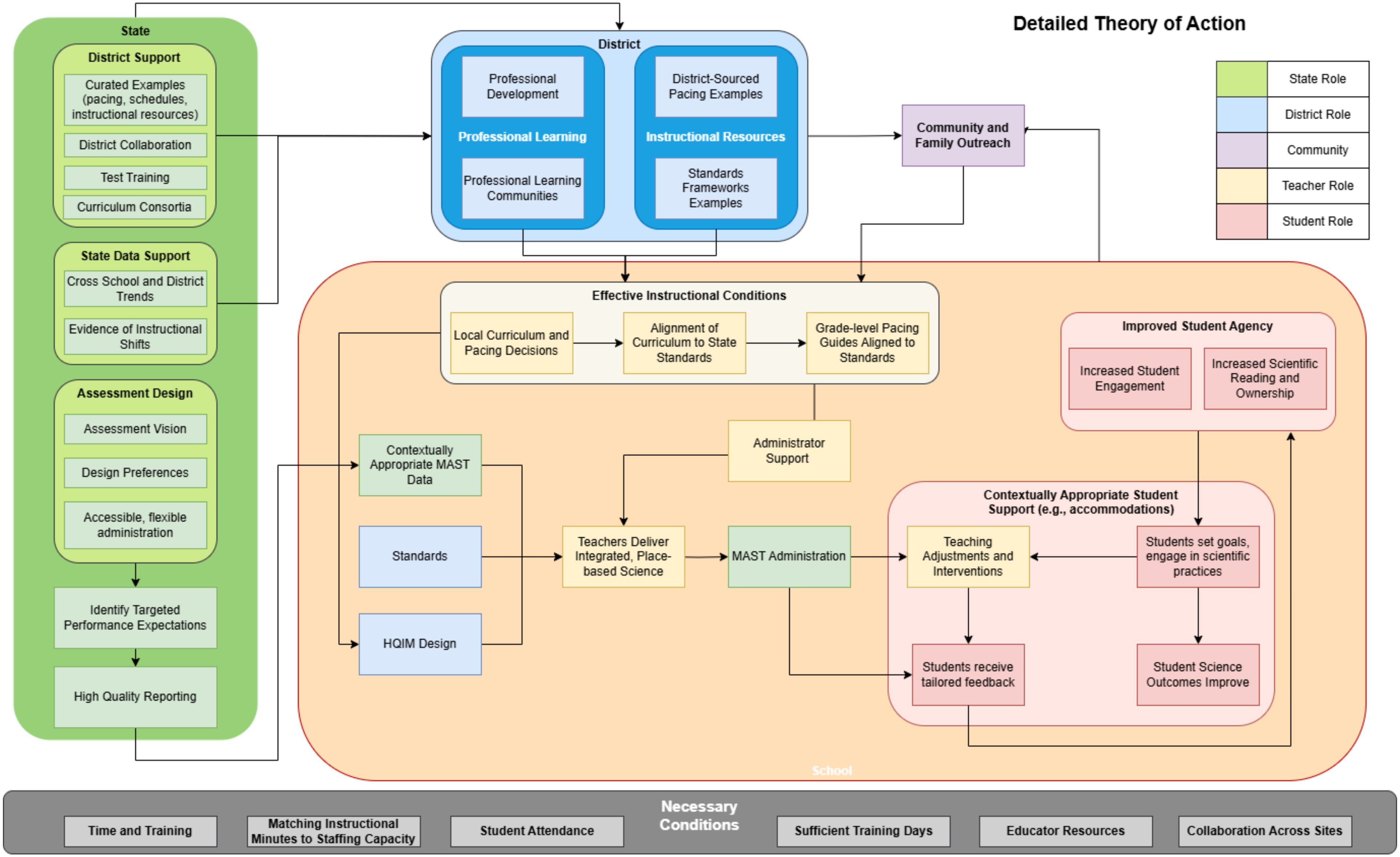
This initial view of the Detailed Theory of Action portrays the interconnected roles and responsibilities of the state, districts, schools, teachers, and students in Montana's MAST Science system. It highlights how each part of the system contributes to creating effective instructional conditions and improving student agency in science.

At the left, the state's responsibilities are clearly defined: designing the assessment around targeted performance expectations, generating high-quality reporting, and supporting districts with data use and implementation resources. This state infrastructure enables districts to organize professional learning and develop high-quality curriculum and instructional resources aligned to state standards.

Districts also play a key role in facilitating community and family outreach, which supports broader understanding and buy-in for the assessment system. These combined efforts flow into the creation of effective instructional conditions at the school level, where teachers—supported by administrators—deliver integrated, place-based science instruction informed by MAST data. Within classrooms, students engage in scientific practices, set personal learning goals, and receive tailored feedback. Contextually appropriate accommodations and interventions are made to ensure all students can engage meaningfully with the content. The feedback loop between student engagement, tailored support, and instructional adjustments is central to the theory of action, driving continuous improvement in science outcomes and fostering student ownership of learning.

This visualization reinforces the idea that student success emerges from coherent coordination across the system, with clearly delineated support roles and shared responsibility for implementation.

Figure 3. Comprehensive Theory of Action for Montana's Science Assessment System.



This comprehensive visualization represents the most detailed iteration of the Combined Theory of Action for Montana's MAST Science assessment system. It maps out a coherent system of roles, supports, and instructional conditions, highlighting how state, district, community, school, and student responsibilities align to support improved science learning outcomes across Montana.

On the left, the state's foundational responsibilities are shown in three distinct areas: Assessment Design, State Data Support, and District Support. These components include articulating a clear assessment vision, supporting accessible and flexible administration, identifying targeted performance expectations, and offering high-quality reporting. In parallel, the state also facilitates data interpretation, tracks evidence of instructional shifts, and strengthens district capacity by providing curated resources, test training, and support for curriculum consortia. These supports enable districts to focus on professional learning and instructional resource development. District responsibilities include organizing professional learning communities, offering targeted professional development, and sharing district-sourced pacing examples and frameworks aligned to standards. A critical component is the emphasis on family and community outreach, demonstrating the importance of communicating with parents and guardians as part of implementation and use.

These upstream supports result in the creation of "Effective Instructional Conditions" within schools. This area reflects the direct alignment of local curriculum and pacing with state standards, including grade-level guides and collaborative decisions agreed upon by teachers and administrators. Teachers, supported by administrator leadership, deliver integrated, place-based instruction rooted in both state expectations and local relevance.

The Detailed TOA also emphasizes the role of MAST data as a tool for tailoring instruction. Teachers use contextualized MAST data to make timely adjustments and to implement targeted interventions, supported by clear structures for accommodations and tailored feedback. This instructional response cycle supports students in setting goals, engaging with science meaningfully, and developing greater ownership over their learning.

At the top right, the end goals are made explicit: increased student engagement, greater scientific reasoning, and improved science learning outcomes. These gains are achieved not by a single entity but through the careful orchestration of supports, systems, and conditions across the educational continuum.

Finally, the bottom bar highlights "Necessary Conditions" required for successful implementation. These include adequate time and training, alignment of staffing and instructional time, sufficient training days, and collaborative structures—especially important for Montana's small and rural districts. By explicitly naming these prerequisites, the Detailed TOA acknowledges implementation challenges while pointing to the enabling conditions required for sustained success.

Together, this Detailed Theory of Action offers a blueprint for operationalizing MAST Science. It articulates not only what is needed at each level of the system, but also how those pieces fit together to create a coherent, responsive, and improvement-oriented assessment system.

Recommendations for Implementation

This section translates the Theory of Action into a practical set of recommendations for launching and sustaining the MAST Science assessment system. These recommendations are intended to support pilot planning, statewide rollout, and ongoing system improvement.
The theory of action calls for the concerted effort of all of those involved in both the teaching and assessment of science in Montana. This means that the scope of the theory of action, while anchored in statewide summative assessment, extends far beyond the MAST science assessment itself. Given this, the recommendations for implementation also extend outside of the assessment program to include the effective use and interpretation of assessment results.

1. **Establish a Design and Implementation Advisory Group.** To ensure the assessment system remains responsive to emerging needs and grounded in local expertise, a standing advisory group should be created. This group can provide feedback and help course-correct during the pilot and scale-up phases. It should be diverse in composition to ensure representation of both small and large districts, rural and urban communities, and varied instructional capacities. This group would:
   - Guide design revisions and pilot implementation.
   - Represent the needs of rural, small, and under-resourced districts.
   - Monitor system coherence and feedback loops across implementation phases.

2. **Prioritize Standards and Claims to Guide Assessment Development.** A focused and clearly articulated set of performance expectations will provide structure for task development, scoring, and reporting. Anchoring assessments to the prioritized outcomes developed by the Task Force will help ensure alignment and coherence across instructional materials, assessments, and student reporting. This will:
   - Ensure that assessment tasks are aligned with meaningful learning outcomes.
   - Support the development of performance level descriptors (PLDs) and task models.
   - Help districts sequence instruction and support vertically aligned learning.

3. **Build Capacity for Local Implementation.** To bring the Theory of Action to life, those in schools and districts will need support, especially given limitations in time and resources. In a number of places in Montana, a school *is* a district, with educators simultaneously fulfilling multiple roles. Given this, capacity building efforts should work

across organizational boundaries to pool resources and maximize the amount of support available along the lines explained within the *Examples of Roles and Supports at Each Level* subsection. At a high level, these supports may involve:

- Professional learning on data use, instructional alignment, and the purposes of through-year assessment.
- Guidance for curriculum consortia serving small or rural districts.
- Sample pacing guides and exemplar units to ease integration with instruction.

The following recommendations focus on how to approach capacity building based on the relationships and partnerships between groups and organizations.

- The **state can:**
    - Share high quality practice by convening interested schools and districts virtually so that these practices can be shared directly, and/or by curating materials from these practices and making them available online through, for example, a learning hub website. Especially important in this work is ensuring that the varying capacities of schools and districts are considered (e.g., an approach for one room school houses and an approach for larger districts with multiple schools).
    - Partner or otherwise collaborate with regional curricula consortia on various issues, including on how the forthcoming assessment program aligns to these instructional materials.
    - Provide materials, training, and eventually assessment results to support school and district leaders in providing instructional support, coaching and mentorship to teachers, in situations where there are different people in these roles.
    - Collaborate directly with teachers and leaders, especially to identify the specific ways in which the score reports and supporting materials are meant to inform practice, and then to develop these reports and materials.
- **Schools** and **districts can**:
    - Connect teachers to one another across schools and districts to form professional learning communities, especially in cases where teachers do not have access to such communities within their own organizations.
    - Support teachers in making connections with parents, potentially by providing state communication resources or developing their own.

These kinds of capacity-building efforts should be ongoing and contextually responsive, addressing both instructional and technical needs.

4. **Pilot and Evaluate with Transparency and Flexibility.** The pilot phase will offer critical insights into the practicality and effectiveness of the system design. A structured, transparent evaluation approach will foster trust while ensuring that lessons learned

inform future iterations. Flexibility will allow the system to adapt to varied district capacities and priorities. Recommendations include:

- Establishing clear criteria for evaluating pilot success (e.g., use of data by teachers, alignment with instructional practices).
- Collecting feedback from both high-capacity and low-capacity districts.
- Sharing pilot findings with educators and families to build trust in the system.

5. **Design Reporting Tools That Support Interpretation and Action.** Score reports should be more than a summary—they should be tools that support reflection, dialogue, and instructional decision-making. Effective reporting will require a clear visual design, relevant content, and support to help educators, families, and students interpret and act on the data. Reports should:
   - Be timely and user-friendly for educators, families, and students.
   - Include both individual and aggregate results that inform instruction and decision-making.
   - Offer plain-language interpretations of claims and suggestions for next steps.

6. **Communicate and Engage Continuously.** Sustained family and community engagement is essential for system legitimacy and student support. Communication should not be an afterthought; it should be embedded throughout the implementation process, with dedicated time, tools, and personnel focused on outreach and responsiveness. Recommendations to support community use include:
   - Co-designing communication materials with district leaders and family engagement coordinators.
   - Providing multi-modal supports (e.g., videos, guides, training) for using MAST results.
   - Creating space for feedback from end-users throughout rollout and beyond.

7. **Develop Implementation Guide for Schools and Districts.** The piloting and operational administration of the MAST science assessment program should operate in parallel with a set of activities aimed at improving and supporting science instruction in the state. These activities can be supported by a phased implementation guide that explains key milestones towards the kinds of teaching and learning that support the outcomes envisioned by the task force.

The Role of Reporting

Reporting is a cornerstone of any assessment system. It is the mechanism through which the results of student learning are communicated and acted upon. In the context of the MAST Science assessment system, reporting plays a vital role in bridging the gap between assessment and instruction. Effective reporting not only makes assessment data accessible but also translates those data into insights that can inform teaching practices, student support strategies, and family engagement. For Montana's through-year science system to be successful, reporting must be timely, clear, and actionable across audiences.

The Montana Task Force emphasized that reporting is not simply the last step in the assessment cycle but an integral part of a continuous improvement loop. If the MAST Science system is to support place-based learning, differentiated instruction, and equitable access to science learning opportunities, its reporting tools must provide information that educators and families can understand and use. This means moving beyond raw scores or summative labels to reports that help teachers identify strengths and needs, help students track their own growth, and help families engage in meaningful conversations about science learning.

To fulfill this role, reporting for MAST Science should adhere to several key principles that emerged throughout Task Force conversations. These are presented in the table below.

Table 5. Reporting Principles and Descriptions

Reporting Principle	Description
Clarity and Accessibility	● Reports should use plain language and avoid technical jargon wherever possible. ● Visual representations (e.g., charts, growth arrows, or comparison bands) should support understanding for educators, families, and students. ● Translations and digital accessibility options should be made available to ensure equitable access for all families.
Timeliness	● Through-year testing allows for multiple data points across the year. Reports should be available soon after each administration window so that the results can inform instruction in real-time. ● Teachers should receive interim results while instruction is still in progress, enabling them to adjust instruction and provide targeted support.
Actionability	● Reports should connect to instructional next steps, such as standards-aligned strategies or recommended resources. ● They should include clear descriptions of what students know and can do, linked to the performance expectations assessed. ● For families, reports should include suggestions for supporting learning at home or questions to ask during teacher conferences.
Multi-level Reporting	● Reporting should be available at multiple levels (student, classroom, school, district) to support a range of users. ● Aggregated reports should help school and district leaders identify trends, guide professional development, and allocate resources.

Alignment with Claims and Prioritized Outcomes	● Reports must reflect the core claims and prioritized outcomes identified by the Task Force. ● This includes emphasizing scientific reasoning, data literacy, and real-world application, rather than rote content knowledge.
Integration with Communication Tools and Structures	● Reporting should not exist in isolation; it should be part of a larger communication strategy that includes family engagement materials, educator training, and public transparency tools. ● Clear explanations of how results are generated, what they mean, and how they should (and should not) be used can build trust in the system.

Ultimately, effective reporting transforms data into insight. It allows teachers to better support their students, empowers students to take ownership of their learning, and equips families with the tools to advocate for their children. In Montana's MAST Science system, high-quality reporting is not a technical afterthought. Reporting is a foundational design principle and a driver of system equity, transparency, and instructional relevance.

Appendix A: Role-Based Considerations for Approaching the Theory of Action

At the State Level

The success of Montana's science assessment system depends on thoughtful policy leadership and strategic oversight. At the state level, leaders are responsible for setting a vision, aligning policy, and sustaining long-term commitment to the goals laid out in the theory of action. This section provides guidance to help state board members and legislators understand how their decisions shape system coherence, fairness, and support for districts and schools.

State Board of Education Overview

Question	What to Consider
How should I read this document?	Focus on the overarching goals, vision, and coherence of the assessment system with statewide education priorities.
What can I do to help support this?	Endorse policies that ensure sustainability, alignment, and transparency across systems.
What part of the theory of action should I focus on?	System-wide coherence and accountability structures.
What questions should I ask next?	How will we evaluate impact over time? How does this align with long-term state goals?

State Board of Education. As a State Board of Education member, this report should be read as a strategic roadmap for advancing the quality, coherence, and equity of science education across Montana. It outlines a vision for a system that is grounded in student learning, anchored in shared goals, and designed for sustainable impact. Your role is pivotal in maintaining the visibility and legitimacy of this work over time.

To help implement this, consider advocating for and adopting policies that sustain a coherent assessment system over time, including guidelines for accountability, implementation timelines, and educator supports. Your endorsement of this work can create the policy infrastructure needed to reinforce local efforts and secure cross-administration continuity. Be prepared to champion a balanced approach to innovation that keeps student learning at the center while providing districts and educators with the time and resources they need.

In the theory of action, your primary area of focus should be system-wide coherence. This includes the structures and supports that ensure each level of the system—from the state to the classroom—is aligned around shared learning goals, expectations, and reporting. You are uniquely positioned to ensure that policies reflect the core principles of the TOA and that system components (e.g., curriculum, instruction, professional learning, accountability) work in harmony.

Key questions to ask include: How will we evaluate the long-term effectiveness and equity of the system? What accountability structures must evolve to support this new approach? What data and evidence will we use to make course corrections or expand this work statewide? How can we ensure coherence across other subject areas and system initiatives?

State Legislature Overview

Question	**What to Consider**
How should I read this document?	Understand the rationale and goals behind the redesign and its implications for educational equity and local innovation.
What can I do to help support this?	Support legislation and funding for pilot implementation, educator training, and system development.
What part of the theory of action should I focus on?	The relationship between assessment practices and long-term student outcomes.
What questions should I ask next?	What will the return on investment be for local communities? How does this align with Montana values?

State Legislators. As a state legislator, this report serves as a lens into the thinking behind Montana's effort to modernize science education. You should read it with an eye toward understanding how the proposed assessment system advances the goals of equity, innovation, and student readiness for future academic and career pathways. The document outlines how assessment can act as both a learning tool and a policy lever to elevate instructional quality across diverse communities.

You can help implement this vision by supporting legislation that provides stable funding for the pilot and scale-up of the through-year assessment. You can emphasize that the science assessment embodies Montana's constitutional requirement and duly enacted Indian Education for All (IEFA) policy. This includes ensuring resources are available for educator training, technology upgrades, and the development of reporting systems that return useful information to families and schools. Legislators can also foster cross-agency partnerships and create incentives for local innovation that align with the system's goals.

In the theory of action, your area of focus should be on the connection between assessment and

Montana's broader workforce development and civic engagement strategies. The TOA emphasizes the importance of preparing students not only for tests but also for life, ensuring they can think critically, solve real-world problems, and actively participate in their communities.

Key questions to ask include: What will the return on investment be for our communities? How does this new system increase opportunities for students in rural and underserved areas? What evidence will we use to determine if the system is working as intended? How can state policy accelerate and protect these efforts over time? How will we monitor equity of access and outcomes?

At the District Level

District leaders serve as the bridge between state vision and school-level implementation. Superintendents and district staff play a critical role in communicating goals, aligning resources, supporting schools, and ensuring that the assessment system serves instructional improvement.

This section outlines how district-level leaders can engage with the theory of action and coordinate coherent implementation across schools.

District Superintendents Overview

Question	What to Consider
How should I read this document?	As a strategic roadmap for shaping instructional priorities and communicating expectations to schools and families.
What can I do to help support this?	Champion district-wide participation, align curriculum and assessment efforts, and engage families.
What part of the theory of action should I focus on?	District-level supports and communication loops.
What questions should I ask next?	How can we tailor implementation to meet our local needs? What support do our staff need?

District Superintendents. As a district superintendent, this document provides a strategic overview of the direction Montana is taking in science assessment. Read it as both a policy guide and a tool for aligning district efforts with statewide priorities. Pay close attention to the implementation guidance and examples of support structures, as they directly affect your local responsibilities.

To support implementation, superintendents should lead efforts to communicate the vision and purpose of the system within their district. This includes facilitating cross-role collaboration, ensuring schools are prepared for the pilot phase, and advocating for professional learning that enables teachers to interpret and use assessment information effectively. Superintendents can help connect this work with existing district improvement plans and ensure that resources—both human and technical—are allocated strategically.

Within the theory of action, superintendents' key focus is on building district-level coherence and capacity. This involves supporting school leaders and staff as they implement the system, using data to inform continuous improvement, and modeling the culture of instructional support and collaboration. Superintendents are vital in shaping how the system comes to life across schools.

Key questions to ask include: How can we align this system with our local goals and capacity? What structures will we need to support teachers and principals in using the new system? How can we ensure small and rural schools in our district receive equal support? What role should district leadership play in monitoring and refining implementation?

District Staff Members (Assessment, Curriculum, etc.) Overview

Question	What to Consider
How should I read this document?	As a technical guide to help align instructional resources, pacing guides, and professional development.
What can I do to help support this?	Build internal coherence between district tools and the assessment system.
What part of the theory of action should I focus on?	Instructional alignment and capacity-building supports.
What questions should I ask next?	How can we use this system to improve teaching and learning?

District Staff Members (Assessment, Curriculum, Instructional Leads). As district staff, this document should be read as a framework to guide your coordination of curriculum, assessment, and instructional supports. The report provides key insights into how assessment can be embedded in instruction and how reporting structures should empower local decision-making.

You can support implementation by mapping district curriculum and pacing guides to the new assessment structure, co-developing professional learning opportunities with building leaders, and acting as liaisons between state-level updates and school-level needs. Your work ensures coherence between what students are taught, how they're assessed, and how results are used. In the theory of action, your primary area of focus is instructional alignment and the utility of assessment tools. You play a vital role in making sure teachers have the right supports, resources, and professional development to make the system effective at the classroom level.

Key questions to ask include: How can we use this system to improve instructional coherence? What tools or guidance do teachers need to act on results? How will we support principals and teacher leaders in interpreting and using assessment data? What processes should we build to sustain and refine this work beyond the pilot year?

At the School Level

Schools are where the assessment system becomes real for students. Principals, teachers, and teacher leaders are on the front lines of implementation, translating the vision into daily classroom experiences. This section describes how school-based educators can use the theory of action to improve teaching and learning, foster collaboration, and ensure assessment practices are meaningful and student-centered.

School Principals Overview

Question	**What to Consider**
How should I read this document?	As a framework for supporting teachers and aligning school goals with statewide outcomes.
What can I do to help support this?	Allocate time and resources for teacher collaboration and assessment use.
What part of the theory of action should I focus on?	School-level structures that enable effective instructional practices.
What questions should I ask next?	How can I create a school culture that embraces assessment for learning?

School Principals. As a school principal, this report serves as both a vision-setting document and a practical guide to implementing a new science assessment system at your school. Read it with an eye toward how assessment design connects to your role in instructional leadership, teacher support, and community engagement.

To support implementation, create structures for professional learning and collaboration that help teachers interpret assessment data and use it to inform instruction. Lead efforts to embed the assessment system into school improvement plans and ensure alignment between instructional priorities and assessment goals, including alignment for IEFA. Consider establishing regular data team meetings, creating a school-based implementation team, or aligning professional development days to focus on science assessment and instructional alignment.

Your focus in the theory of action is school-level capacity-building. This means fostering a culture of evidence-informed teaching, encouraging professional conversations around student learning, and ensuring the assessment experience reflects the values of your school community. Principals are uniquely positioned to serve as communicators between district expectations and teacher realities. They can shape a collaborative culture where assessment is used to inform—not punish—teaching and learning.

Key questions to ask include: What support do my teachers need to feel confident in using this system? How can I ensure consistency of implementation across grade levels? How can we communicate the purpose and value of the assessment system to families and students? What professional learning routines can we build into our school year to keep this work sustainable?

School Teachers Overview

Question	**What to Consider**
How should I read this document?	As a guide to understand what students should know and be able to do, and how assessments will support your instruction.
What can I do to help support this?	Participate in pilots, use results to inform instruction, and provide feedback.
What part of the theory of action should I focus on?	Classroom-level processes and instructional feedback loops.
What questions should I ask next?	How will this support what I'm already doing in the classroom?

School Teachers. This document is your guide to how science assessment in Montana is evolving to be more supportive of learning. Read it as both a vision and a source of practical insight into what the system aims to do for you and your students.

You can support implementation by sharing feedback during the pilot, participating in professional learning communities, teacher partnerships, vertical and horizontal teams, and using the assessment data to adjust your instruction in real time. Teachers are also encouraged to surface implementation barriers and suggest improvements that support both students and colleagues. For example, you might bring up how well the timing or format of the assessment fits with your unit pacing or suggest ways for integrating assessment results into class discussion, student reflection, or parent-teacher conferences.

In the theory of action, your key focus is the day-to-day use of assessment information to improve student learning. You are central to ensuring that the system supports growth, encourages inquiry, authentically embeds IEFA, and reflects the diverse needs of your learners. Teachers have the clearest lens on what works at the classroom level and can serve as important advocates for making the system more relevant and responsive.

Key questions to ask include: How can I use results from this system to guide instruction and student goal-setting? What professional learning will help me use these tools more effectively? How can I collaborate with peers to ensure shared understanding and expectations? What opportunities can I create for students to engage in self-assessment or reflect on their science learning?

School Curriculum Leaders and Teacher Leaders Overview

Question	**What to Consider**
How should I read this document?	As a leadership tool to support alignment between curriculum and assessment across grades and teams.
What can I do to help support this?	Lead training and conversations about how performance expectations show up in local curriculum.
What part of the theory of action should I focus on?	Translating system goals into practical classroom supports.
What questions should I ask next?	How do we scaffold teacher understanding and use of the system?

Curriculum Leaders and Teacher Leaders. As curriculum and teacher leaders, this report provides strategic context for your work in designing instruction, facilitating collaboration, and mentoring peers. Read it as both a planning tool and an opportunity to lead system-wide improvements in instructional practice.

Support implementation by facilitating professional development, aligning instructional materials with assessment tasks, and creating opportunities for teachers to engage in collaborative planning. You are key connectors between vision and practice. You might organize model lesson studies, co-design formative assessment tasks with teachers, or curate resources that connect science-aligned instruction with assessment anchors.

In the theory of action, your focus is ensuring that instructional support reflects the vision of deeper learning, IEFA integration, and real-world application. You also play a key role in helping teachers interpret and act on student data in meaningful ways. This includes developing routines for data interpretation and helping teachers translate results into actionable instructional moves.

Key questions to ask include: How can I support my colleagues in navigating this change? What common planning time or protocols are needed to ensure effective collaboration? How do our current resources align with the expectations of this system? How can we build internal capacity to sustain assessment literacy across staff?

Families and Communities

Family engagement is essential to making the assessment system work for students. When families understand the goals of the system and how it supports learning, they can better advocate for and support their children. This section helps parents and guardians understand their role in the vision and how they can stay informed and involved in their child's science education.

Parents and Guardians Overview

Question	**What to Consider**
How should I read this document?	As a window into what your child is learning and how schools are supporting science learning across grades.
What can I do to help support this?	Stay engaged with teachers, use reports to ask questions, and support learning at home.
What part of the theory of action should I focus on?	Student experience and family engagement components.
What questions should I ask next?	How can I understand my child's progress? What can I do to support science learning at home?

Parents and Guardians. Read this document as a way to understand how Montana is working to make science assessment more relevant, meaningful, and supportive of your child's education. The goal is not just to test students, but to provide information that helps everyone—families included—support learning.

You can help by staying engaged in your child's science learning, asking questions about what they're learning and how they're being assessed, and advocating for opportunities for all students to explore science in hands-on, real-world ways. For example, you can ask your child's teacher how science is being taught throughout the year, what assessments your child will take and why, and how the results will be shared and used to support learning.

Your role in the theory of action is as a partner. The system works best when families understand its goals and participate in conversations about learning, progress, and support. Families can play an active role in reinforcing science thinking at home and helping students see themselves as capable science learners.

Key questions to ask include: How will I receive information about my child's learning? What can I do at home to support science thinking and inquiry? How do I interpret results from the assessment system to understand my child's growth and needs? Who can I contact at my child's school if I have questions or want to learn more?