Standard-Setting Technical Report

for the

Maine Science Assessment

Grades 5, 8, and High School

Prepared for Public Posting by New Meridian Corporation and the Maine Department of Education

Table of Contents

Introduction

The Maine Science Assessment is a component of the Maine Comprehensive Assessment System aligned to the Maine Learning Results (MLRs), for which the Next Generation Science Standards (NGSS) serve as the foundation. The Maine Science Assessment is administered at grades 5, 8, and 3rd year of high school. The purpose of the computer-delivered assessment is to measure students' science knowledge, skills, and abilities.

The Maine Science Assessment was developed by New Meridian based on the Maine DOE-approved test blueprints aligned to the NGSS. The assessment measures key Disciplinary Core Ideas (DCIs), Science and Engineering Practices (SEPs), and Cross-Cutting Concepts (CCCs). The first operational administration was in May 2022.

The Maine DOE approved policy achievement level descriptors (ALDs) for four levels—*Well Below State Expectations*, *Below State Expectations*, *At State Expectations*, and *Above State Expectations* (see Appendix A). To report Maine Science Assessment results, cut scores (or threshold scores) would need to be set by the Maine DOE. Maine educators were convened to (a) review the operational assessments administered in May 2022, (b) describe the science knowledge and skills necessary for students in grades 5, 8, and the third year of high school to be placed into one of four achievement levels, and (c) recommend cut scores to the Maine DOE. To develop cut score recommendations for each grade level, New Meridian conducted a standard-setting workshop in Augusta, Maine, for the Maine DOE on July 26–28, 2022.

Purpose and Background

The purpose of standard setting for the Maine Science Assessment was to gather recommendations from Maine educators for the Maine DOE to review and consider when adopting cut scores for reporting results. For each grade-level assessment, there are four performance levels. A cut score defines the beginning of a higher level of performance or achievement. Therefore, the standard-setting panelists defined threshold students and made judgments for cut scores marking the boundary between *Well Below* and *Below State Expectations*, *Below* and *At State Expectations*, and *At* and *Above State Expectations*.

A review of the standard-setting literature supports the need for attention to best practices (Hambleton, Pitoniak, & Copella 2012; Tannenbaum & Katz, 2013), which include the following:

- A careful selection of panelists to represent varying perspectives
- Sufficient time devoted to developing a common understanding of the assessment domain
- Adequate training of panelists
- Development of a description of student performance at each threshold

- Multiple rounds of judgments
- The inclusion of data, where appropriate, to inform judgments

The approach used in this study, the Bookmark standard-setting method (e.g., Karatonis & Sireci, 2006; Lewis, Mitzel, Mercado, & Schulz, 2012; Mitzel, Lewis, Patz, & Green, 2001), adheres to these guidelines.

Methodology

New Meridian used the modified Bookmark standard-setting procedure to set the three cut scores within each Maine Science Assessment grade that differentiate the achievement expectations for these four levels:

- Well Below State Expectations
- Below State Expectations
- At State Expectations
- Above State Expectations

Policy Achievement Level Descriptors (ALDs) were developed by the Maine DOE and envision the knowledge, skills, and abilities (KSAs) of students in each of the achievement levels (described in Appendix A).

Modified Bookmark Standard-Setting Procedure

The Bookmark standard-setting procedure (Lewis, Mitzel & Green, 1996) was developed in 1996 and fully documented in *Setting Performance Standards: A Guide to Establishing and Evaluating Performance Standards on Tests* (Cizek & Bunch, 2007). The iterative procedure is a complete set of activities intended to produce cut scores based on the judgments of subject matter experts (i.e., educators serving as standard-setting panelists). The Bookmark standard-setting procedure was developed to provide a less complex method for panelists to determine a cut score (Mitzel et al., 2001) and may be the most commonly used procedure for establishing cut scores for large-scale K–12 assessments (Baldwin, Margolis, Clauser, Mee & Winward, 2019).

The key activities for the Maine Science Assessment standard setting included the following:

- Subject-matter experts reading and answering tasks for the Maine Science Assessment
- Group discussion of ALDs and students at the “just barely” threshold
- Subject-matter experts’ examination of the ordered item booklet (see below) and group discussion of items
- Training on the Bookmark standard-setting process

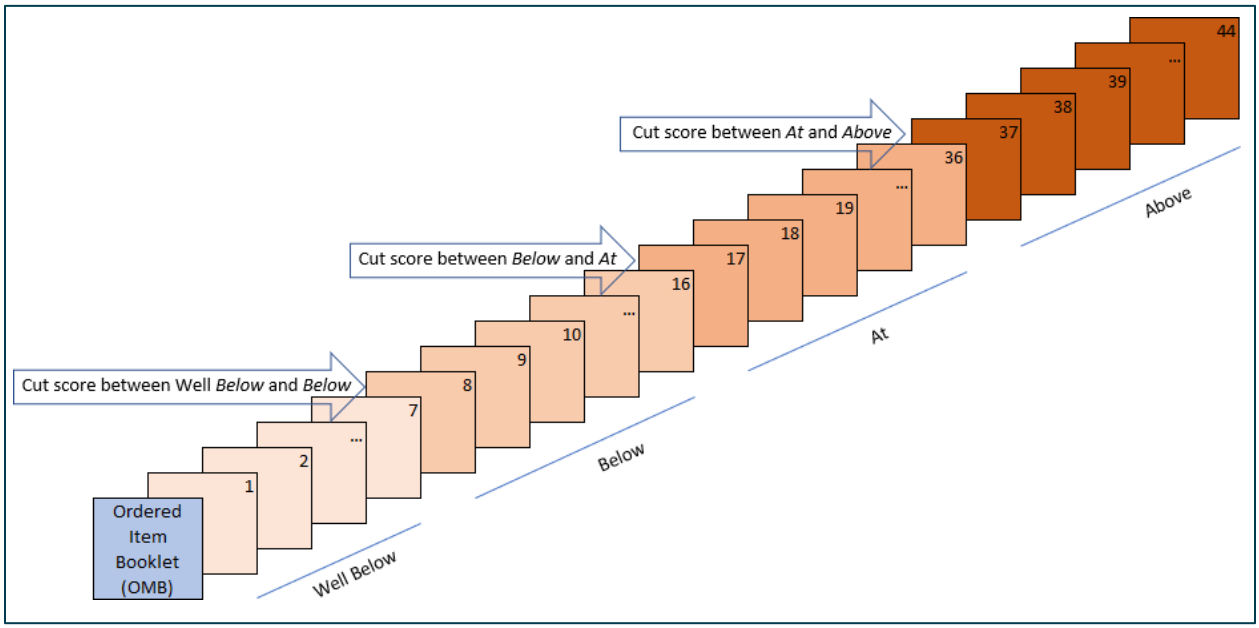
- Individual panelist's bookmark placements (Round 1 ratings)
- Group discussion of Round 1 results and impact data
- Individual panelist's bookmark placements (Round 2 ratings)
- Group discussion of Round 2 results and impact data
- Individual panelist's bookmark placements (Final ratings)
- Group discussion of Final results and impact data

The Ordered Item Booklet (OIB)

The Ordered Item Booklet (OIB) is integral to the modified Bookmark standard-setting procedure. In this method, test items (i.e., questions) were ordered in the OIB from easiest to most difficult based on their statistics estimated from operational administration data (May 2022). Each page in the OIB represents one item or score point. Questions worth 2 points are represented in the OIB with one page for each possible point value: 1 point (partial credit) and 2 points (full credit).

Figure 1 shows an example of an OIB for an assessment with four reported achievement levels. Items at the beginning of the OIB are aligned with KSAs that *Well Below State Expectations* students can typically demonstrate based on the ALDs. Items in the middle of the OIB are aligned to *Below State Expectation* students followed by items aligned to At *State Expectations* students. Items at the high end are aligned to *Above State Expectations* students. The points in the OIB that locate the boundaries between the achievement levels are the cut scores.

Example OIB for a Multi-level Passing Scores



Response Probability

New Meridian used a 0.67 response probability criterion (RP67) for the Maine Science Assessment standard setting. The RP value is the probability for answering an item correctly for a given level of student ability that is located on the score scale. New Meridian used RP67 for each item to find the location on the score scale (i.e, ability level) where a student would have 0.67 probability for answering the item correctly; and that location is where the item is mapped to on the score scale. For instance, if RP67 for a given item corresponds with the cut score for At State Expectations, and a given student has an ability score at that cut score, then the student will have a response probability of 0.67 (or a 67% chance) for answering the item correctly. All items were mapped to the score scale using the response probability criterion, which resulted in the ordering of items along the score scale that served as the basis for the ordering of items in the OIB.

RP67 is the most common response probability criterion for the Bookmark standard-setting procedure. RP67 separates items that a threshold student has and has not mastered and is a relatively easy value for panelists to understand (Mitzel et al., 2001; Karantonis & Sireci, 2006).

Individual Panelist Bookmark Placement

A key element to the Bookmark procedure is the reconciliation of the ALDs and the threshold student's demonstrated KSAs. During standard setting, panelists proceed through the OIB one page at a time and ask themselves whether they believe the just barely, or threshold, student would have a 0.67 probability of answering the item correctly (or obtaining the given score point). If the answer is "Yes," the panelist proceeds to the next page. The page on which the panelist answers "No" becomes the bookmark between the adjacent performance levels (e.g., between *Well Below* and *Below State Expectations*). This process is used for all three cut-score placements.

Presentation of Results and Iterative Group Discussion

Following each of the three rounds, New Meridian collected the panelists' bookmark placements (provided as a page number) and compute the median, minimum, and maximum placement for the panel. New Meridian also converted these results to the corresponding RP67 values.

New Meridian presented these results and the corresponding impact data, which is the percentage of students at each of the four achievement levels based on the proposed bookmark placements. The panelists then discussed the subset of items within the range of individually bookmarked pages and considered the ALDs and the KSAs a threshold student requires. Following the Round 1 and the Round 2 group discussions, panelists had the opportunity to update their bookmark location based on the shared discussion. At the culmination of the Final round, the committee's median RP67 value was used to calculate the associated cut scores with respect to the reported score scale and the raw score scale.

Standard-Setting Panels

The centerpiece of the standard-setting workshop was the recruitment and selection of subject-matter experts (or panelists). Best practice assumes that the selected panelists reflect the educators that instruct students who will take the Maine Science Assessment because they have a thorough understanding of the assessed content.

The intent was for the workshop panel for each grade to include ten panelists recruited by the Maine DOE. The ultimate size of each panel was six panelists for Grade 5, five panelists for Grade 8, and eight panelists for High School.

Training and Workshop Activities

Workshop Materials

New Meridian set up a secure Smartsheet dashboard through which panelists accessed workshop tools. The page included the following items:

- Event agenda
- Link to Maine Science Assessment forms
- Reference materials (e.g., content standards)
- Item map table
- Training and Bookmark placement forms
- Links to reflection and evaluation forms

The following materials were provided as hard copies during the workshop:

- Test Blueprints
- Achievement Level Descriptors
- Ordered Item Booklet

These workshop materials were assigned to each panelist and were not allowed outside of the workshop meeting area. Secure materials were collected by New Meridian staff prior to the workshop being adjourned each day.

Facilitators

New Meridian conducted the standard-setting workshop, and New Meridian staff served as facilitators and provided information and resources but did not contribute to the cut-score recommendations during the workshop.

The New Meridian team included facilitators and science content leads, as well as program management staff, to facilitate the use of Smartsheet and all planning during the workshop.

Panelist Training

In the week prior to the workshop, all panelists were asked to complete a set of pre-workshop activities to familiarize themselves with the assessment and the information to be used throughout the standard-setting workshop.

The pre-workshop activities included the following:

- **Completing a nondisclosure agreement** certifying that the panelists will keep all workshop materials and discussions secure
- **Reviewing tutorials** that introduce the testing environment, the item types comprising the assessments, and the various tools students have access to when completing the assessment
- **Reviewing reference materials** to be used during the workshop
- **Completing a survey** with information about the panelists' current teaching position, teaching experience, and demographic information

The standard-setting workshop took place in Augusta, Maine on July 26–28, 2022. Except for introductory comments on Day 1, panelists worked in separate rooms during the workshop. The following table provides the agenda for the workshop.

Standard-Setting Workshop Agenda

Time	Activity
Day 1 (July 26, 2022)	
11:00 AM	Registration/Check In Opens
11:30 AM	Lunch
12:00 PM	Opening Session: Welcome from Maine DOE and Dashboard Training
12:45 PM	Review of Reference Materials and Tutorials
1:45 PM	Break
2:00 PM	Sit the Maine Science Assessment Exam
3:30 PM	Wrap-up for Day 1
3:45 PM	Adjourn Day 1
Day 2 (July 27, 2022)	
8:00 AM	Registration/Check In Opens
8:15 AM	Day 1 Recap
8:30 AM	Discuss the Achievement Level Descriptors (ALDs) and Define the Threshold Students
10:30 AM	Break
10:45 AM	Discuss the Threshold Student Across Groups
11:45 AM	Discuss the Ordered Item Booklet (OIB)
12:45 PM	Lunch
1:15 PM	Discuss the Ordered Item Booklet (OIB) continued

2:15 PM	Bookmark Placement Training and Practice
2:45 PM	Round 1 Bookmark Placement (Three Cuts)
3:45 PM	Wrap-up for Day 2
4:00 PM	Adjourn Day 2
Day 3 (July 28, 2022)	
8:00 AM	Registration/Check In Opens
8:15 AM	Day 2 Recap
8:30 AM	Present Round 1 Results & Round 2 Directions
8:45 AM	Discuss Round 1 Results
10:15 AM	Break
10:30 AM	Round 2 Bookmark Placement (Three Cuts)
11:30 AM	Lunch
12:00 PM	Present Round 2 Results & Final Round Directions
12:15 PM	Discuss Round 2 Results
1:15 PM	Final Bookmark Placement (Three Cuts)
2:15 PM	Break
2:30 PM	Presentation of Final Recommendations
2:45 PM	Workshop Evaluation
3:00 PM	Wrap-up Day 3
3:15 PM	Adjourn Day 3

Discussion of the ALDs and the Threshold Students

In smaller groups within their grade-level rooms, panelists first discussed the ALDs and the threshold students who just barely met the requirements for each achievement level. Panelists were provided with Maine DOE-approved policy ALDs to gain an understanding of the high-level expectations for students within the four achievement levels.

Next, panelists reviewed threshold ALDs to summarize the expectations for students who have just enough KSAs to be considered in that achievement level. Panelists were encouraged to imagine a hypothetical threshold student to represent this determination and to engage in structured discussions about the KSAs they expect to be demonstrated by the threshold student.

Finally, the panelists reconvened as a panel in their grade-level room to discuss the ALDs for the four performance levels and the differences between them, considering the overall level of rigor implied by the policy ALDs. To focus on the line of demarcation between two adjacent performance levels, panelists were asked to discuss the KSAs that separate students into the levels. Panelists sketched a “picture” of the threshold students that could be referenced throughout the workshop. The expectations for the threshold students were based on the content standards and the policy ALDs. Appendix B contains the sketches of the threshold

students for each grade level.

Discussion of the OIBs and Item Maps

Panelists reviewed the OIB item by item and considered what each item measures and why it was more difficult than the items preceding it. The item map included additional information about each item to describe the item type, DCI and SEP being measured, etc. Panelists were instructed to take notes about the KSAs required to answer the items correctly.

Bookmark Placement Training

New Meridian explained and illustrated what the bookmarks mean and what panelists should consider in placing them. The training described how cut-score recommendations could be represented by bookmarks and explained that all items preceding the bookmark contain the KSAs that a student who is just barely in the level is expected to know.

Panelists were trained to provide a content-based rationale for their bookmark that referred to the alignment between the KSAs in the ALDs and those in the items before the bookmark.

Following training, panelists completed a short quiz to verify their understanding of the procedure. New Meridian provided each correct answer to the quiz with an explanation to ensure all panelists understood the process before proceeding to Round 1 judgments.

Bookmark Placement Rounds and Discussions

During each bookmark placement round, panelists independently placed their three bookmarks (one for each cut score) and recorded their placements and content-based rationale for the placement in a Smartsheet form. These content-based rationales were to solely be used for panelists' reference during their panel's discussion before the next round.

Following each round of bookmark placements, New Meridian calculated the bookmark recommendation for the panel. Panelists were presented with a summary of their round recommendations, including the median, minimum, and maximum for bookmark placement. Grade-level impact data was also shared.

Panelists discussed the rationale behind their bookmark placement, focusing on the items in the OIB between the lowest and highest bookmarks they placed for each cut in the round. Panelists referred to their OIBs, item maps, ALDs, and the content standards throughout the discussions.

Following Round 1 and Round 2 discussions, panelists were reminded that they may maintain their bookmark placement from the previous round or move the bookmarks, provided the panelists supply a content-based rationale for any change.

Final Recommendations

Following the final placement, the panelists were presented the final cut-score recommendations and impact data based on the median of the final round of bookmark placements.

Workshop Results

Recommended Threshold Score Results

The tables below display the median bookmark threshold scores after each round for Grades 5, 8, and High School. The median was calculated for the overall panel. The tables show how panelists moved the bookmarks across rounds. Lower numbers represent bookmark placements earlier in the OIB, indicating a threshold score on a less difficult item that translates to a lower threshold score. Higher numbers translate to a higher threshold score; a higher threshold score means that more is required for a student to be included in the level. In general, the threshold scores were consistent between Rounds 2 and 3 except for the *At/Above* threshold for Grade 8. (Differences between Rounds 1 and 2 were observed for all thresholds at Grades 5 and 8.)

Median Bookmarks by Round: Grade 5

Threshold	Round 1	Round 2	Final
Well Below/Below	7	8	8
Below/At	17	21	20
At/Above	38	37	37

Median Bookmarks by Round: Grade 8

Threshold	Round 1	Round 2	Final
Well Below/Below	6	5	5
Below/At	18	13	13
At/Above	34	21	31

Median Bookmarks by Round: High School

Threshold	Round 1	Round 2	Final
Well Below/Below	6	6	6
Below/At	23	23	23
At/Above	43	42	43

Impact Data

Panelists viewed the impact data associated with the median bookmark judgments for each round.

The impact data in the following table shows the percentage of students who would be placed in each achievement level based on the Final round median threshold score recommendations. These impact data are based on the May 2022 operational administration of the Maine Science Assessment and may differ from the percent of students in each level in future test administrations.

Impact Data based on Final Round Recommendations by Grade

Achievement Level	Grade 5	Grade 8	High School
Well Below State Expectations	44.39%	33.37%	41.51%
Below State Expectations	31.31%	18.67%	20.12%
At State Expectations	20.03%	44.01%	32.24%
Above State Expectations	4.27%	3.96%	6.13%

Post-study Analyses

Following the final round, New Meridian analyzed the recommended cut scores with the empirical data while investigating potential adjustments to present to the Maine DOE with the associated impact data. Adjustments to committee-recommended cut scores derived from test-centered approaches to standard setting are commonly used to ensure the cut scores are practical while still reflecting the committee's content-based recommendations.

Conditional Standard Error of Measurement (CSEM)

The conditional standard error of measurement (CSEM; Kolen & Brennan, 2014) quantifies the amount of statistical error associated with any point on the assessment scale. If a student were to take a Maine Science Assessment form multiple times, New Meridian expects the hypothetical student's test score to fall within a range of ±1 CSEM about two-thirds of the time. If the difference between two test scores is less than one CSEM, it is often considered difficult to describe the difference as significant. CSEM values are frequently used to assess cut-score precision and to adjust cut scores after standard-setting.

Results and Recommendations

Grade 5 Recommendations

For grade 5, the cuts recommended by the panel would result in approximately 44%, 31%, 20% and 4% of students falling into the corresponding achievement levels. Potential

adjustments and the matching impact data are also provided in the table below. Adjusting the cut scores by subtracting one or two CSEMs will result in more students falling into the higher achievement levels, while adding one or two CEMs will result in fewer students at the higher achievement levels.

Grade 5 Impact Data (Percentages) for Recommended Cut Scores and Adjustments

Achievement Level	-2CSEM	-1 CSEM	Recom. Cut	+1CSEM	+2CSEM
Well Below State Expectations	23.66%	31.79%	44.39%	57.14%	72.38%
Below State Expectations	29.28%	33.30%	31.31%	28.01%	20.16%
At State Expectations	29.18%	25.14%	20.03%	13.58%	7.07%
Above State Expectations	17.88%	9.76%	4.27%	1.27%	0.39%

Maine DOE accepted the panel's recommended cuts.

Grade 8 Recommendations

For grade 8, the panel's recommended cuts would result in approximately 33%, 19%, 44% and 4% of students falling into the corresponding achievement levels. Potential adjustments and the matching impact data are also provided in the table below.

Grade 8 Impact Data (Percentages) for Recommended Cut Scores and Adjustments

Achievement Level	-2CSEM	-1 CSEM	Recom. Cut	+1CSEM	+2CSEM
Well Below State Expectations	11.99%	24.31%	33.37%	47.24%	66.41%
Below State Expectations	12.31%	13.72%	18.67%	19.17%	12.66%
At State Expectations	58.42%	52.81%	44.01%	32.27%	20.52%
Above State Expectations	17.27%	9.17%	3.96%	1.32%	0.40%

Maine DOE accepted the panel's recommended cuts.

High School Recommendations

For High School, the panel's recommended cuts would result in approximately 42%, 20%, 32% and 6% of students falling into the corresponding achievement levels. Potential adjustments and the matching impact data are also provided in the table below.

High School Impact Data (Percentages) for Recommended Cut Scores and Adjustments

Achievement Level	-2CSEM	-1 CSEM	Recom. Cut	+1CSEM	+2CSEM
Well Below State Expectations	22.32%	30.26%	41.51%	50.16%	58.80%
Below State Expectations	19.19%	19.90%	20.12%	20.13%	19.84%
At State Expectations	37.13%	37.86%	32.24%	26.19%	19.79%
Above State Expectations	21.36%	11.99%	6.13%	3.52%	1.57%

Maine DOE accepted the panel's recommended cuts.

References

Baldwin, P., M. J. Margolis, M.J., Clauser, B.E., Mee, J., Winward, M. (2019). The choice of response probability in bookmark standard setting: An experimental study. *Educational Measurement: Issues and Practice, 39*(1), 37–44.

Cizek, G. & Bunch, M. (2007). Standard setting: A guide to establishing performance standards on tests. Thousand Oaks, CA: Sage Publications.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37–46.

Hambleton, R. (2001). Setting performance standards on educational assessments and criteria for evaluating the process. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods, and perspectives*. Mahwah, ME: Lawrence J. Earlbaum.

Hambleton, R. K., Pitoniak, M. J., & Copella, J. M. (2012). Essential steps in setting performance standards on educational tests and strategies for assessing the reliability of results. *Setting performance standards: Foundations, methods, and innovations*, *22, 47– 76.*

Lewis, D., Mitzel, H., & Green, D. (1996, June). Standard setting: A bookmark approach. In D. Green (Chair), *IRT-based standard setting procedures utilizing behavioral anchoring*. Symposium conducted at the Council of Chief State School Officers National Conference on Large-Scale Assessment, Phoenix, AZ.

Lewis, D. M., Mitzel, H. C., Mercado, R. L., & Schulz, E. M. (2012). The bookmark standard setting procedure. *Setting performance standards: Foundations, methods, and innovations*, *22*, 225-54.

Karantonis, A., & Sireci, S. (2006). The Bookmark Standard-Setting Method: A Literature Review. *Educational Measurement, 25*, 1, 4–12.

Kolen, M. & Brennan, R. (2014). *Test equating, scaling, and linking: Methods and practices*. (3rd ed.) New York, NY: Springer-Verlag.

Livingston, S. A., & Lewis, C. (1995). Estimating the consistency and accuracy of classifications based on test scores. *Journal of Educational Measurement, 32*(2), 179–197.

Mitzel, H., Lewis, D., Patz, R., & Green, D. (2001). The Bookmark procedure: Psychological perspectives. In G. Cizek (Ed.), *Setting performance standards: Concepts, methods and perspectives* (pp. 249–281). Mahwah, ME: Erlbaum.

Tannenbaum, R. J., & Katz, I. R. (2013). Standard setting. In K. F. Geisinger, B. A. Bracken, J. F. Carlson, J.-I. C. Hansen, N. R. Kuncel, S. P. Reise, & M. C. Rodriguez (Eds.), *APA handbook of testing and assessment in psychology, Vol. 3. Testing and assessment in school psychology and education* (pp. 455–477). American Psychological Association.

Appendix A. Policy Achievement Level Descriptions

Level 1 Well Below State Expectations	Level 2 Below State Expectations	Level 3 At State Expectations	Level 4 Above State Expectations
The student's work demonstrates a minimal understanding of essential concepts in science. The student's responses demonstrate minimal ability to solve problems. Explanations are illogical, incomplete, or missing connections among central ideas. There are multiple inaccuracies.	The student's work demonstrates an incomplete understanding of essential concepts in science and inconsistent connections among central ideas. The student's responses demonstrate some ability to analyze and solve problems, but the quality of responses is inconsistent. Explanation of concepts may be incomplete or unclear.	The student's work demonstrates an adequate understanding of essential concepts in science, including the ability to make connections among central ideas. The student's responses demonstrate the ability to analyze and solve routine problems and explain central concepts with sufficient clarity and accuracy to demonstrate general understanding.	The student's work demonstrates a thorough understanding of essential concepts in science, including the ability to make multiple connections among central ideas. The student's responses demonstrate the ability to synthesize information, analyze and solve difficult problems, and explain complex concepts using evidence and proper terminology to support and communicate logical conclusions.

Appendix B. Threshold ALDs

New Meridian staff transcribed the notes created and used by standard-setting committee members to define threshold ALDs. Transcriptions were not edited to standardize grammar or conventions.

Grade 5

Life Science

Well Below/Below

- Understands there is a food chain/web/ but is unclear of the order
- Understand that plants & animals have different parts/body
- Struggle to describe how the parts function

Below/At

- Understands the process & order of plant/prey/pred.
- May struggle w/ decomposers/fungi come into play.
- Demonstrate different functions
- Struggle to identify their purpose

At/Above

- Understands that a change has occurred w/ a new species but is unsure of what those effects will be.
- Understand that living things are made up of cells.
- Struggles w/ cell functions

Physical Science

Well Below/Below

- Liquid 1 Liquid 2 = total
- Volume is obj1 + obj 2 = combined
- Describe why a balloon expands
- Simple classifications based on visible properties

Below/At

- Describe matter in a popped or filled balloon
- Draw a simple diagram of a filled or deflated balloon
- Sometimes you can distinguish between physical vs. chemical change, i.e., vinegar and baking soda

At/Above

- Describe the phenomenon of changed particles and temperature
- Chemical reactions, law of conservat [conservation]. Describe phenomenon
- Can they describe conservation of matter.

Earth Science

Well Below/Below

- Label concrete objects
- Use a model to identify similarities/differences
- Recognizes the sun is a star
- Can identify day and night with a diagram or basic model

Below/At

- Can explain big ideas with support of a word bank or model/diagram
- Create a basic model to explain seasons/night & day (patterns)
- Analyze the distance determines the brightness of a star

At/Above

- Can modify a model to support an argument
- Clear, concise, accurate vocabulary in constructive response
- Can mathematical computation to support an argument “What If”
- Analyze the distance & size determine the brightness of a star

Grade 8

Initially, the middle school educators broke into three groups by subject matter expertise, to define threshold ALDs. All then joined together to synthesize the ALDs from the individual subjects. In this section the synthesized ALDs will be presented.

Well Below/Below

- Use model
- Compare observ.
- ID Qualitative Patterns
- Define Vocab
- Read a Graph
- Follow Investigat.
- Distinguish Qual vs. Quant data

Below/Meets

- Develop Model (with support)
- Obs. Connects to Claim
- Explain Variable Relations
- Relations between Vocab
- Interpret a graph
- ID Variables (independent dependent)
- Analyze Data

Meets/Above

- Revise Model
- Synthesize Obs. (multiple pieces)
- Extrapolate and Reason
- Integrated Vocab. Used Correctly
- Create or Analyze a Graph
- Design Experiment
- Derive Meaning from Data

High School

In the high school room, the educators were divided into groups by discipline expertise (life science, physical science, and Earth and space science). Each group came up with descriptors of the 3 threshold students (well below/below, below/at, and at/above state expectations).

Well Below/Below

Life Science Group

CAN	CAN'T
Matter and Energy	
• Ask questions about how plants convert light energy to sugars as a food source	• Develop a model to show connections between photosynthesis and C.R. [cellular respiration]
Interdependent Relationships in Ecosystems	
• Ask questions about populations	• Use model to make connections about other factors that are impacting them (temp, predator, biotic, abiotic)
• Act to increase/decrease #'s [numbers] survive	• Cause & effect
Natural Selection and Evolution	
• Ask questions about how the environment impacts the existence of a species	• Use data
	• Recognize that in any species there is genetic variability that provides some individuals with an advantage

Inheritance and Variation of Traits

- Able to ask questions about why certain phenotypes are more common
- Recognize that there is variability between species
- Explain the molecular basis of the differences between species due to genetics

Structure and Function

- Ask clarifying questions
- Realize that there are different cell types
- Use a model to show interactions of body structures

Earth and Space Science Group

CAN

- Identify trends on graph or 2 graphs
- Make accurate observations of a model and ask questions about why it looks that way.
- Describe qualitative quantities (less/more)

CAN'T

- Describe how 2 or more graphs are related
- Make reasonable claims based on observations of a model (Why do you think" Idk)
- Model differences in quantities such as w/ proportions

Physical Science Group

CAN

- Read and understand data
- Use data to make simple conclusions
- Make a hypothesis
- Conduct a "cookbook procedure" to gather data
- Identify patterns in data
- Ask/evaluate a question that requires limited evidence/data

CAN'T

- Create models
- Link evidence & conclusions

Below/At Expectations

Life Science Group

CAN	CAN'T
Matter and Energy	
• Develop and model to show connection between photosynthesis & CR (matter & energy)	• Use data to validate the model & don't recognize energy is lost to the environment
Interdependent Relationships in Ecosystems	
• Model how environmental changes affect populations of species	• Recognize that certain populations have a carrying capacity using data/graph (math)
Natural Selection and Evolution	
• Use data to recognize variability – some advantageous as the population evolves	• Use/Apply Evidence to show Common Ancestry
Inheritance and Variation of Traits	
• Explain that a trait comes from genetics (DNA, Genom, Etc.)	• Create a model to show that the differences are due to DNA mutations (Environmental Factors)
Structure and Function	
• Use a model to show interactions between body parts	• Develop a model to show how a system is regulated or how it maintains homeostasis. (Feedback loops)

Earth and Space Science Group

CAN	CAN'T
• Observe trends, describe relationship between models	• Incorporate other factors into their explanation. i.e. thinking critically about relationships outside the model & their effects
• When given data, can describe relationships but struggle with creating a model representing that relationship	• Identify data needed to Support their claims

Physical Science Group

Note: *The picture of the list that the Physical Science group produced was too blurry to be transcribed.*

At/Above

Life Science Group

CAN	CAN'T
Matter and Energy	
• Use data to show photosynthesis & CR cycle matter and energy is lost	• Reason that sugar molecules are the basis of carbon compounds but modified
Interdependent Relationships in Ecosystems	
• Use math/data to support the concept of a carrying capacity	• Develop a model expansive of a disturbance on an entire ecosystem
Natural Selection & Evolution	
• Use evidence to show common ancestry	• Make predictions using math/statistics about the distribution of certain phonotypes
Inheritance and Variation of Traits	
• Use a model to connect mutations to the variant	• Use data to predict the genotype from a pedigree of phenotypes

Life Science Group (cont.)

CAN	CAN'T
Structure and Function	
• Develop a model to show how systems are regulated (Homeostasis)	• Use DNA evidence to recognize differential protein production and regulation and gene expression

Earth and Space Science Group

CAN	CAN'T
• Describe relationships between 2 or more models	• Make predictions from the models
• Develop models when given data	• Use those models to explain novel scenarios or apply to other situations
• Explain a phenomenon using Scientific Concepts	• Identify when data is needed to those support concepts

Physical Science Group

CAN	CAN'T
• Develop effective models based on data • Independently develop & conduct experimental procedures to test a question • Support an explanation with multiple pieces of data • Use mathematical models to support a claim	• Evaluate models • Identify limitations of models • Evaluate/revise a procedure to improve data collection