Personalized Alternate Assessment Portfolio

MeCAS Part II

2018–19 TECHNICAL REPORT

© 2019 Cognia, Inc.

TABLE OF CONTENTS

CHAPTER 1 OVERVIEW 1
1.1 PURPOSE OF THIS REPORT 1
1.2 ORGANIZATION OF THIS REPORT 2
CHAPTER 2 CURRENT YEAR UPDATES 3
CHAPTER 3 THE STATE ASSESSMENT SYSTEM 4
3.1 INTRODUCTION 4
3.2 ALTERNATE ASSESSMENT BASED ON ALTERNATE ACHIEVEMENT STANDARDS 5
3.3 THE ALTERNATE ASSESSMENT SYSTEM 5
3.4 PURPOSES OF THE ALTERNATE ASSESSMENT SYSTEM 5
3.5 GENERAL FORMAT AND BACKGROUND 6
CHAPTER 4 THE STUDENTS 8
4.1 PARTICIPATION DECISION PROCESS 8
4.2 SUMMARY OF PARTICIPATION RATES 9
CHAPTER 5 TEST CONTENT 10
5.1 ALTERNATE GRADE LEVEL EXPECTATIONS 10
5.1.1 Levels of Complexity 10
5.1.2 Format of the AGLEs for the PAAP 10
5.2 ACCESS TO THE GENERAL CURRICULUM 12
5.3 ASSESSMENT DESIGN 12
5.4 ASSESSMENT DIMENSIONS 15
5.5 TEST SECURITY 18
CHAPTER 6 TEST DEVELOPMENT 19
6.1 GENERAL PHILOSOPHY 19
6.2 ROLE OF COMMITTEES IN TEST DEVELOPMENT 20
CHAPTER 7 ALIGNMENT 21
7.1 DESCRIPTION OF LINKAGES TO DIFFERENT CONTENT AREAS ACROSS GRADES 21
CHAPTER 8 PAAP ADMINISTRATION TRAINING 22
8.1 STEPS FOR ADMINISTRATIONS 22
8.2 STEPS IN CONSTRUCTING THE PORTFOLIO 23
CHAPTER 9 SCORING 28
9.1 TABLE LEADER AND SCORER RECRUITMENT AND QUALIFICATIONS 28
9.2 TABLE LEADER AND SCORER TRAINING 28
9.3 SCORING PROCESS 29
9.4 FLOW OF MATERIALS 31
9.5 SECURITY 32
9.6 SCORING RUBRIC 32
9.7 CALCULATION OF REPORTED SCORES 35

CHAPTER 10 CLASSICAL ITEM ANALYSIS 36
10.1 DIFFICULTY AND DISCRIMINATION 36
10.2 STRUCTURAL RELATIONSHIP 38
10.3 BIAS/FAIRNESS 39
CHAPTER 11 CHARACTERIZING ERRORS ASSOCIATED WITH TEST SCORES 40
11.1 RELIABILITY 40
11.2 SUBGROUP RELIABILITY 41
11.3 DECISION ACCURACY AND CONSISTENCY 42
11.4 INTERRATER CONSISTENCY 44
CHAPTER 12 COMPARABILITY (SCALING AND EQUATING) 45
12.1 COMPARABILITY OF SCORES ACROSS YEARS 45
12.1.1 Reported Scores 45
12.1.2 Standard Setting 45
12.2 LINKAGES ACROSS GRADES 46
CHAPTER 13 VALIDITY 48
13.1 EVIDENCE BASED ON TEST DEVELOPMENT AND STRUCTURE 48
13.2 OTHER EVIDENCE 49
13.3 FUTURE DIRECTIONS 49
REFERENCES 50
APPENDICES 51
APPENDIX A 2019 ALTERNATE GRADE LEVEL EXPECTATIONS
APPENDIX B PROCESS FOR DETERMINING THE APPROPRIATE AVENUE FOR PARTICIPATION
APPENDIX C 2019 SCORING INSTRUCTIONS
APPENDIX D SCORE OF RECORD
APPENDIX E ITEM-LEVEL CLASSICAL STATISTICS
APPENDIX F ITEM-LEVEL SCORE DISTRIBUTIONS
APPENDIX G SUBGROUP RELIABILITY
APPENDIX H DECISION ACCURACY AND CONSISTENCY RESULTS
APPENDIX I INTERRATER CONSISTENCY
APPENDIX J CUMULATIVE SCORE DISTRIBUTIONS
APPENDIX K ACHIEVEMENT-LEVEL DISTRIBUTIONS
APPENDIX L PROCESSING AND REPORTING BUSINESS REQUIREMENTS

CHAPTER 1 OVERVIEW

This technical report provides an overview of Maine's alternate science assessment, the Maine Educational Assessments (MEA) Alternate Science—Personalized Alternate Assessment Portfolio (PAAP), which is administered to students with significant cognitive disabilities who, because of their unique learning needs, cannot access the MEA Science even with a combination of accommodations. Descriptions of the purpose of the PAAP, the processes utilized to develop and implement the PAAP program, and stakeholder involvement in those processes are included in this report. By comparing the intent of the PAAP with its process and design, the assessment's validity can be evaluated. Stakeholder groups such as the PAAP Advisory Committee, item/task review committees, and content committees helped guide the development and implementation process. Teacher input in the development of the overall PAAP process is described, from the Alternate Grade Level Expectations (AGLEs) design through blueprint/test design, content alignment, task development, task tryout/field-testing, teacher trainings, test administration, scoring, and standard setting.

1.1 PURPOSE OF THIS REPORT

The purpose of this report is to document the technical aspects of the 2018–19 PAAP operational implementation. Science was assessed at grades 5 and 8 and 3rd year high school.

Several technical aspects of the PAAP are described in an effort to contribute to evidence supporting the validity of PAAP score interpretations. Because the interpretations of test scores, not the test itself, are evaluated for validity, this report presents documentation to substantiate intended interpretations (AERA et al., 2014). Each chapter contributes important information to the validity argument by addressing one or more of the following aspects of the PAAP: task development, alignment, administration, scoring, reliability, standard setting, achievement levels, and reporting.

Standards for Educational and Psychological Testing (AERA et al., 2014) provides a framework for describing sources of evidence that should be considered when constructing an argument for assessment validity. These sources include evidence in five general areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. Although each of these sources may speak to a different aspect of validity, the sources are not distinct types of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

1.2 ORGANIZATION OF THIS REPORT

This report is organized based on the conceptual flow of the PAAP's yearlong process, which includes blueprint design/development, task development, administration, scoring, reporting of scores, technical characteristics, and validity. The appendices contain supporting documentation.

CHAPTER 2 CURRENT YEAR UPDATES

There were minimal changes from previous administrations to the assessment of science through the PAAP for the 2018–19 assessment year. Formal reporting was removed from this contract starting with the 2015–16 administration and is handled by the Maine DOE's reporting vendor. Measured Progress merged with Advance Education in January of 2019 and renamed their company in July of 2019 to Cognia. Cognia delivers final data files to the Maine DOE for this assessment. Test Security Agreements for Test Administrators were included in the 2017 handbook and remain a permanent part of the PAAP Handbook.

CHAPTER 3 THE STATE ASSESSMENT SYSTEM

The Maine Comprehensive Assessment System (MeCAS) is a statewide instructionally supportive assessment system that complies with the federal requirements of the Every Student Succeeds Act (ESSA) and the Individuals with Disabilities Education Improvement Act (IDEA) of 2004, as amended. These acts, along with state regulations, require that all students, including those with disabilities, participate in state-mandated assessments in grades 3–8 and 3rd year of high school, and they are intended to hold schools accountable for the academic performance of students. Those assessments include:

- Cognia eMPowerME, which assesses mathematics and English language arts (ELA)/literacy, at grades 3–8
- Maine SAT, via the College Board, which assesses mathematics and ELA/Literacy at 3rd year high school
- the MEA Science, which assesses science at grades 5 and 8 and 3rd year of high school
- Personalized Alternate Assessment Portfolio (PAAP), which alternately assesses science, for a small number of students with the most significant cognitive disabilities who are unable to take part in the general science assessment
- Multi-State Alternate Assessment (MSAA), which assesses students at grades 3–8 and 3rd year high school with the most significant cognitive disabilities who are unable to take part in the mathematics and ELA/literacy assessment

All students participate in the statewide assessment in one of three ways: general assessment, general assessment with accommodations, or alternate assessment, as outlined in the following sections.

3.1 INTRODUCTION

The PAAP, like the MEA Science, is designed to provide a snapshot in time of an individual student's performance. A broader picture will emerge as the student results on the PAAP are reviewed along with results on other formative and summative assessments.

PAAP tasks are provided in the PAAP Task Bank for the science content standard Levels of Complexity (LoCs) as described in the PAAP Alternate Grade Level Expectations (AGLEs) document (see Appendix A). Tasks selected for use in an individual student's PAAP should match the instructional level at which the student is working and be designated within the PAAP AGLEs/Indicators as appropriate for his or her grade level.

The AGLE/Indicators include LoC descriptors that have been reduced in complexity to ensure access to instruction and assessment for all students.

All tasks submitted in a PAAP are corrected (by item), resulting in an overall percentage score for the task. The evidence (student work) included in a 2018–19 PAAP for science must have been generated during the PAAP test administration window: March 1, 2019 –April 30, 2019.

3.2 Alternate Assessment Based on Alternate Achievement Standards

Prior to 2015, 1% of Maine students in grades tested may show academic proficiency through administration of an alternate assessment based on alternate achievement standards. With the implementation of ESSA regulation 114-95 (ll) in 2015, 1% of all students eligible to test in grades 5, 8, and third year high school may be assessed for alternate assessments. The PAAP is designed for students with significant cognitive disabilities that are unable to participate in the general MEA even with the best instruction and appropriate accommodations.

As previously described, the PAAP is designed under the guiding philosophy that alternate achievement standards are built on measurable, targeted skills linked to Maine's *2007 Learning Results* for science, and represent student performance at a lower level of breadth, depth, and complexity than that found in the general assessment.

3.3 The Alternate Assessment System

Given the legislative context within which the entire statewide assessment system sits, the PAAP is, as a part of the overall MeCAS, governed by the same laws and rules that govern general assessment. Federal legislation, including IDEA of 2004, and ESSA of 2015, requires that students with disabilities have access to the general curriculum, with appropriate accommodations where necessary, and that they be assessed on the same general curriculum standards as all other students. For the small number of students who cannot participate in the general large-scale assessment due to their significant cognitive disabilities, the law also allows—and Maine provides—a statewide alternate assessment based on the AGLEs. Alternate achievement standards are reduced in breadth, depth, and complexity, while maintaining linkage to the same general curriculum standards taught to all children.

3.4 Purposes of the Alternate Assessment System

The PAAP is designed to provide instruction and a meaningful academic assessment experience in science, based on the AGLEs, for those Maine students with the most significant cognitive disabilities.

The portfolio approach captures student progress in academic content over the course of a five-month instructional window, with actual assessment occurring March 1 through April 30, 2019, enabling teachers and others to see evidence of this progress within the context of the instructional program they are providing.

The PAAP is also intended to provide feedback to teachers on student performance, which they can use to improve instruction.

As part of this purpose, the PAAP signals to Maine special education teachers the need to maintain high academic expectations for their students and high standards in the delivery of their instructional programs. Students receive greater learning opportunities throughout their academic careers because of strictly followed test blueprints and teacher trainings that encourage educators to move PAAP students to higher Levels of Complexity.

The PAAP ensures that all Maine students are appropriately included in state and federal assessment requirements and provides instructional improvement that reveals what students know and are able to do. This system aims to meet the highest technical standards possible while best serving the students participating in the assessment.

3.5 General Format and Background

AGLE Entries submitted in a PAAP must comprise four components:

- an Entry Slip that serves as the organizer for all student work related to a single content standard
- the required number of Task Descriptions designed to help the user understand the expectations of an individual task, how the task was administered, the prior knowledge required to perform the task, and the alignment to the specific standard and performance indicator being measured
- the required quantity of student work to serve as evidence of student performance (see Appendix A)
- a Task Summary page summarizing the Level of Accuracy and Level of Assistance

Forms for the Entry Slips and Task Descriptions have been common since 2003. From 2002 to 2004, only teacher-developed tasks were used in PAAPs. Teacher training on the PAAP process included tools to ensure alignment to the rubrics, sufficiency of evidence, and clarity for scorers. During the 2003 and 2004 scoring sessions, scorers were asked to identify tasks they saw as “exemplary”—those tasks that, clearly aligned, provided evidence of a pattern of performance and could be reliably scored. Those exemplar tasks were then reviewed by a group of teachers brought together in the summer of 2004. Members of that group made suggestions for revisions as necessary and eliminated tasks that did not meet the criteria outlined for the review process. The tasks approved by that group served as the basis for the early development of tasks to be included in an online PAAP Task Bank. A few tasks, based on the exemplars and finalized in form by Maine DOE staff, were posted online for optional use in 2004. The number of Task Bank items was expanded in 2004–05 to allow teachers to create an entire 2005 PAAP, including reading, mathematics, and science, without using teacher-developed tasks. The use of teacher-developed tasks was still permitted, however. At

each stage of this development evolution, final tasks were reviewed by members of the PAAP Work Collaborative or the PAAP Advisory Group.

The use of teacher-developed tasks was no longer permitted for the 2006–07 school year. Due to teacher time involved and the variations in the skill levels among teachers for developing tasks, the Maine DOE collaborated with Measured Progress (now Cognia) on the development of new tasks. The first set of tasks produced by Measured Progress (now Cognia)was developed during 2004–05 for use in 2005–06. A second set of tasks was developed in 2005–06 for use in 2006–07. The purpose of this development was to populate an expanded version of the PAAP Task Bank for reading and mathematics.

Teachers completed a Specific Task Feedback Form to provide Measured Progress (now Cognia) and the Maine DOE with guidance on further development and quality assurance of the tasks. Based on the feedback from teachers, all the first-round PAAP tryout tasks were revised by Measured Progress (now Cognia) and the Maine DOE. A second round of development was completed in the summer of 2006 that focused on reading, writing, and science tasks.

In 2007–08, the Task Bank became password protected and was provided solely for the use of Maine teachers developing PAAPs for their students. Because the PAAP is for students with the most significant cognitive disabilities within the Maine school population, the PAAP rubrics were revised to contain only rubric levels 1 and 2.

The 2009–10 assessment program began to move toward a required test blueprint by grade and content area. In developing the blueprint for the PAAP, care was taken to make the progression of tasks parallel to the progression of the general NECAP assessment in all content areas. Teachers were no longer allowed to freely select which AGLEs to assess. Because the Task Bank was not fully populated, teachers were asked to familiarize themselves with the test blueprints for all content areas but to implement the test blueprint for reading only, as reading was the only content area fully populated in the Task Bank. Teachers were not penalized if they did not follow the test blueprint for reading during that assessment year.

Beginning in 2010–11 and continuing through 2013–14, the program provided a fully populated Task Bank for all content areas and enforced the required grade-level test blueprint provided to teachers in 2009–10. Teachers were no longer allowed to select AGLEs for assessment outside of the grade-level blueprint. Instead, the teachers were required to administer the AGLE Entry requirements for each content area. Starting in 2014–15 the PAAP program required teachers to administer the science AGLEs only.

CHAPTER 4 THE STUDENTS

In effective learning environments, instruction and assessment should always be linked. High-quality assessment practices provide information on which to base ongoing development of instruction that is responsive to student needs. In alternate assessment, models of learning and subsequently the linkages between curriculum, instruction, and assessment are deeply impacted by the characteristics of the students themselves. Knowing who these students are and how they learn is critical to the design and development of effective instruction and assessment. In Maine, each PAAP is individualized so that each student's learning needs can be met with instruction that effectively promotes academic growth. The carefully designed common structure underlying the development of every PAAP provides a basis for comparison of performance patterns across students. The structure of the PAAP assessment illustrates both student performance and the student program. In effect, this assessment prioritizes observation of the dynamic links between models of student learning, curriculum, and instruction, and relates them to actual student outcomes. The design of the portfolio is based on the belief that those particular assessment events will allow students to demonstrate their understanding in a given domain, given a particular view of learning that takes into account important individual student differences.

4.1 PARTICIPATION DECISION PROCESS

Students eligible for the 2018–19 alternate science assessment included students with the most significant cognitive disabilities in grades 5, 8 and third year high school. These students need assessments that are individualized and flexible as well as integrated with daily instruction, resulting in student work that provides evidence of what these students are capable of doing. The PAAP was developed as the mode of participation in state assessments for these students. A participation flow chart is provided in the PAAP Handbook and is shown in Appendix B.

All students considered for alternate assessment were reviewed individually by their individualized education program (IEP) team prior to the time of assessment to determine the appropriate avenue of participation, allowing sufficient time for administration of the alternate assessment. This team was to include at least one of the student's teachers, the school's principal or other administrator, the parent(s)/guardian(s), related service personnel, and the student, whenever possible. If it was not possible for the parent/guardian and student to attend the meeting, they were consulted regarding the committee's recommendations. The materials suggested for use at the meeting included (1) the Participating Guidelines(a copy of which is included in Appendix B), (2) the student profile, (3) the approved state assessment accommodations list for the general MEA Science, (4) samples of the student's work, and (5) MEA Science released items to which the samples of the student's work could be compared. The recommendation for a student to take an alternate assessment must be documented in the student's IEP.

4.2 SUMMARY OF PARTICIPATION RATES

Table 4-1 shows a summary of participation in the 2018–19 Maine PAAP by demographic category for science.

Table 4-1. 2018–19 PAAP: Summary of Participation by Demographic Category—Science

Description	*Tested*	
	Number	*Percent*
All Students	385	100.00
Male	248	64.25
Female	137	35.49
Gender Not Reported	0	0.00
Hispanic or Latino	10	2.59
American Indian or Alaskan Native	1	0.52
Asian	11	2.85
Black or African American	21	5.44
Native Hawaiian or Pacific Islander	1	0.26
White (Non-Hispanic)	328	84.97
Two or More Races	12	3.11
No Primary Race/Ethnicity Reported	1	0.26
Currently Receiving EL[1] services	22	5.70
Former EL[1] Student—Monitoring Year 1	0	0.00
Former EL[1] Student—Monitoring Year 2	0	0.00
Former EL[1] Student—Monitoring Year 3	2	0.52
Former EL[1] Student—Monitoring Year 4	2	0.52
EL[1]: All Other Students	324	83.94
Students with an IEP[2]	385	99.74
IEP[2]: All Other Students	0	0.00
Economically Disadvantaged Students	224	60.88
SES[3]: All Other Students	151	39.12
Migrant Students	0	0.00
Migrant: All Other Students	385	100.00
Students Receiving Title 1 Services	0	0.00
Title 1: All Other Students	385	100.00
Plan 504	0	0.00
Plan 504: All Other Students	385	99.22

[1] EL = English Learner
[2] IEP = Individualized Education Program
[3] SES = Socio-Economic Status

CHAPTER 5 TEST CONTENT

Designed specifically for students with significant cognitive disabilities, the PAAP is a portfolio-based test that is aligned with Maine's Alternate Grade Level Expectations (AGLEs). The content of this assessment has been reduced in depth and breadth but remains focused on the AGLEs, which have been linked from the MEA (*2007 Learning Results*) science standards.

5.1 ALTERNATE GRADE LEVEL EXPECTATIONS

The student work included in the PAAP is based on Maine's AGLEs, which are designed for planning and implementing Maine's science alternate assessment. The PAAP measures progress toward the defined AGLEs by allowing students to produce evidence of their knowledge and skills at a specific point in time. It also assesses students at the same grade levels as the MEA Science assessment. The instructional window for the PAAP runs for much of the academic year—from the first week of December through the last week of April. This window provides opportunities for instruction to be embedded in the student's daily work throughout the school year and then be assessed from March 1 through April using PAAP tasks from an online Task Bank.

5.1.1 Levels of Complexity

Maine's AGLEs provide a common basis for the planning and assessment of standards-based instruction and assessment in a system that allows students to work on the AGLE/Indicators, Level of Complexity (LoC) descriptors, and tasks best suited to their individual needs. All tasks submitted in a student's PAAP must be selected and downloaded from the secure, online Task Bank (https://profile.measuredprogress.org/paap/login.aspx). To establish consistency, only items from the Task Bank may be administered.

All tasks within the Task Bank are aligned with Maine's AGLE/Indicator LoCs 1–8. Students working above the LoCs should participate in the standard Maine state assessment at their grade-level placement with appropriate accommodations.

5.1.2 Format of the AGLEs for the PAAP

Maine's AGLEs were formatted by content area, AGLE/Indicators, and LoC descriptors. The Task Bank was made fully operational for the 2012–13 school year. However, because of the implementation of the required grade-level test blueprint, not all LoCs within each AGLE were required for assessment purposes.

Those LoCs that were not required for assessment purposes had the content taken out of the LoC in the AGLE document and were placed in a supplemental document called the Extended Learning AGLEs for teachers to access for instructional purposes.

Figure 5-1 is an example of the science AGLE/Indicator D1.

Figure 5-1. 2018–19 PAAP: Sample Science AGLE/Indicator—D1

Maine's Accountability Standards, Chapter 131 **Science AGLE/Indicator — D1**

The Physical Setting – Universe and Solar System

Student understands the universal nature of matter, energy, force, and motion, and identifies how these relationships are exhibited in Earth Systems, in the solar system, and throughout the universe by:

Level of Complexity 1 (Grades 5, 8, and 3rd Year HS)	Level of Complexity 2 (Grades 5, 8, and 3rd Year HS)	Level of Complexity 3 (Grades 5, 8, and 3rd Year HS)	Level of Complexity 4 (Grades 5, 8, and 3rd Year HS)
describing or otherwise demonstrating understanding of the positions or apparent motions of different objects in our solar system and what these objects look like from Earth by...			
doing the following: • identifying night and day.	**doing both of the following:** • identifying pictures of night and day, **AND** • identifying the Sun and Earth's Moon.	**doing the following:** • identifying the position of the sun at different times by drawing or otherwise describing the movement of the Sun across the sky.	**doing both of the following:** • identifying the position of the sun at different times by drawing or otherwise describing the movement of the Sun across the sky, **AND** • drawing or identifying different phases of the Moon.

Level of Complexity 5 (Grade 8 and 3rd Year HS)	Level of Complexity 6 (Grade 8 and 3rd Year HS)	Level of Complexity 7 (3rd Year HS)	Level of Complexity 8 (3rd Year HS)
explaining the movements and describing the location, composition, and characteristics of our solar system and universe, including planets, the Sun, and galaxies, by...			
doing the following: • identifying the Sun, Earth's Moon, and planet(s) on a given diagram and listing some planets.	**doing the following:** • listing the four inner planets and placing them in order relative to the Sun.	**doing both of the following:** • describing the relative locations of the Sun, Earth, Earth's Moon, and planets and identifying their orbits; **AND** • identifying the role that gravity plays in forming stars, planets, and the solar system.	**doing both of the following:** • identifying the role that gravity plays in forming stars, planets, and the solar system, **AND** • identifying the role that gravity plays in keeping moons in position around planets, and planets around stars.

The header at the top of the sample AGLE page in Figure 5-1 identifies the AGLE as Maine's Accountability Standards, Chapter 131, Maine's *2007 Learning Results* to which this material is aligned. Directly opposite this, on the right side of the field, the corresponding PAAP identifier is situated: Science AGLE/Indicator—D1.

The student expectations for each AGLE—that is, what is being expected of the student in order to demonstrate proficiency as defined in Maine's *2007 Learning Results* for science—are presented in italics below Maine's Accountability Standards, Chapter 131 AGLE. For example, using Figure 5-1 above, the expectations of the student are that he or she "…understands the universal nature of matter, energy, force, and motion, and identifies how these relationships are exhibited in Earth Systems, in the solar system, and throughout the universe…"

Exactly how the student demonstrates understanding of the universal nature of matter, energy, force, and motion, and identifies how these relationships are exhibited in Earth Systems, in the solar system, and

throughout the universe, is detailed in the LoC descriptor table immediately following the student expectations. For example, as shown in Figure 5-1, the student demonstrates understanding of the universal nature of matter, energy, force, and motion by:

> Level of Complexity 1: identifying night and day.
>
> Level of Complexity 2: identifying pictures of night and day, and identifying the Sun and Earth's Moon.
>
> Level of Complexity 3: identifying the position of the Sun at different times by drawing or otherwise describing the movement of the Sun across the sky.

And so on, up to and including LoC 8.

LoCs are ranged 1–8, and each LoC is accompanied by the grade levels for which participation at that LoC is appropriate.

5.2 Access to the General Curriculum

In an effort to document the extent to which students are being exposed to the general curriculum, as required by the Individuals with Disabilities Education Improvement Act of 2004 (IDEA), the achievement standards take into account student access to the general curriculum. The targeted skills taught are connected to the general curriculum standards but are presented in activities that reflect a reduced level of breadth, depth, or complexity. Examples of these targeted skills are found in the online Task Bank by AGLE/Indicator and LoC. Standards-based activities are those learning activities that have outcomes connected to achieving a curriculum framework standard. Activities are evaluated by linkage to grade-level activities. For example, if students in the general education classroom are learning about the solar system and the universe, then the alternately assessed students might be working on activities involving identifying the time of day or the phases of the moon. This activity would be linked to the science standard. Evidence would show application across multiple activities illustrating this skill.

5.3 Assessment Design

Maine's AGLE document was designed to be a bridge to the general curriculum for all students with significant cognitive disabilities who are unable to participate in the general assessment. The individualized education program (IEP) team determines if the student's achievement level on daily work indicates that he or she can participate in the MEA Science through standard administration or administration with accommodation. If the student cannot, the IEP team plans and implements the MEA Alternate Science, or PAAP, for which the student's skills match the PAAP AGLE/Indicators for his or her grade level.

Maine moved to a mandatory PAAP blueprint for 2013–14 (Figure 5-2) that required specific AGLE/Indicators to be assessed at each tested grade level, ensuring that all students had the opportunity to develop an understanding of concepts included in each AGLE/Indicator at the same time as their general

assessment peers. As the blueprint was developed, the design team focused on each content area to ensure that the progression of tasks would parallel the progression in the general assessment. 2013-14 was the last year reading and math was accessed through PAAP. Science only testing has occurred since the 2014-15 assessment year. The final blueprint (Figure 5-2) was reviewed by personnel at the Maine DOE, content specialists at Measured Progress (now Cognia), Maine stakeholders, the PAAP Advisory Committee, and the Technical Advisory Committee. The 2018–19 PAAP Grade-Level Assessments table, as seen in Figure 5-2, outlines the grades and content area assessed through the PAAP.

Figure 5-2. 2018–19 PAAP: PAAP Blueprint

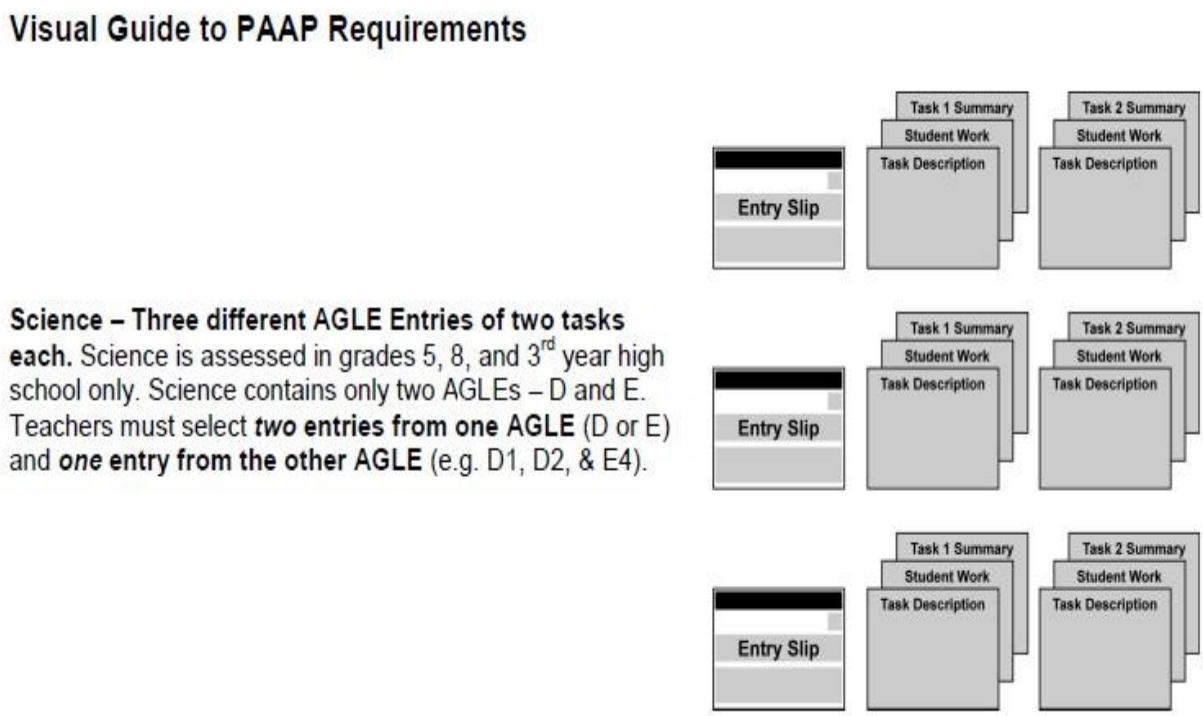
PAAP Blueprint
Required AGLE/Indicators by Content Area

Grade Level	Science
5	D1, D2, E2
8	D4, E3, E4
High School	D3, E1, E5

The PAAP requires four basic components to each AGLE Entry that is assessed: the AGLE Entry Slip, Task Description pages, student work template pages, and Task Summary pages. Science requires three AGLE Entries, each containing two tasks.

Figure 5-3 outlines the PAAP requirements as explained above for the 2018–19 PAAP.

Figure 5-3. 2018–19 PAAP: Visual Guide to PAAP Requirements



As stated in the Chapter 1 Overview, the 2018–19 PAAP was the alternate to the 2018–19 MEA Science. The PAAP assesses two AGLEs: D, the physical setting (D1–D4), and E, the living environment (E1–E5). AGLE D, the physical setting, contains indicators that encompass the subject matter conventionally referred to as physical, Earth, and space sciences, while E, the living environment, contains indicators related to life science.

Indicators from both the physical setting and the living environment are assessed each year in grades 5 and 8 and 3rd year high school. The focus at the elementary level is on concepts that the student can directly observe, such as the Sun, the Moon, rocks, plants, and animals. Force and motion provide concrete observations at the middle school level for the more abstract concepts of matter and energy that will be addressed in high school. Likewise, cells and heredity/reproduction provide foundations for the more abstract concepts of biodiversity and evolution taught in high school, while the level of abstraction increases for the concepts of matter and energy. These are all high school concepts that are more abstract than the concepts covered at the elementary and middle school levels.

In the living environment, the progression from grade 5 to high school is from an understanding of individual organisms and populations to an understanding of how organisms change over time. In the physical setting, the progression is from an understanding of the macroscopic universe, solar system, and Earth to an understanding of forces and motion in the everyday environment, and progressing in high school to an

understanding of matter and energy at the macroscopic and atomic levels. Each successive grade-level assessment connects to and builds on the science concepts introduced at a lower level.

5.4 ASSESSMENT DIMENSIONS

There are three dimensions on which the PAAP is scored:

- Level of Complexity
- Level of Accuracy
- Level of Assistance

Once the AGLE/Indicator for which the student will submit a PAAP has been determined, the teacher chooses the LoC that is appropriate for inclusion in the student's instructional program. The teacher's role is to consider the student's current level of performance and the possibilities for increasing that level through instruction. Teachers may choose a specific LoC and assess the student after instruction has been given. If the student completes that LoC independently with a high percentage of accuracy, the teacher is trained to instruct and assess the student at the next higher LoC. The teacher would then submit the higher LoC to be scored. The same can be done if the teacher assesses at a higher LoC and the student performs below the teacher's expectations (at a very low Level of Accuracy, zero is acceptable) and the student requires the maximum Level of Assistance. The teacher may then back down the instruction to a lower LoC, reassess, and submit the lower LoC to be scored.

The Level of Accuracy on the student work pages is corrected by the teacher item by item on the student work template page, and then the correct/incorrect scores are transferred to the Task Summary page. Each Level of Accuracy box contains the number of items within the task, "Correct/Incorrect" designation with predetermined point values, and the percent correct data key and box. Figure 5-4 is one example of the Level of Accuracy box on the Task Summary page.

Figure 5-4. 2018–19 PAAP: Level of Accuracy

Level of Accuracy						
Item	**Correct/ Incorrect** (Circle One)	**Item**	**Correct/ Incorrect** (Circle One)	Data Key: C = Correct X = Incorrect		
				1 of 6 = 17%	2 of 6 = 33%	3 of 6 = 50%
				4 of 6 = 67%	5 of 6 = 83%	6 of 6 = 100%
1	C X (1 point)	4	C X (1 point)			
2	C X (1 point)	5	C X (1 point)			
3	C X (1 point)	6	C X (1 point)	% Correct = _____		

Students who participate in the state science assessment through the PAAP may need varying degrees of support to complete the required academic tasks. There are three types of support permissible when administering a PAAP:

1. Accommodations selected from the approved list of standard support mechanisms used for general state assessments
2. Flexibility in the method of presentation and student response included within the PAAP directions for task administration
3. PAAP Levels of Assistance

Accommodations do not alter what the test measures or the comparability of results. When used properly, appropriate test accommodations remove barriers to participation in the assessment and provide students with diverse learning needs an equitable opportunity to demonstrate their knowledge and skills.

Accommodations are changes to the standard timing, setting, presentation, or response. An example of an accommodation would be the teacher reading a science task aloud to a student who has a reading disability. The teacher is not altering what is being measured; instead, the student is given the opportunity to demonstrate what he or she knows and can do by eliminating the roadblock his or her disability might otherwise present to the accurate measurement of science knowledge and skills. Students participating in the PAAP may use any of the accommodations that have been approved for use in state assessments by the Maine DOE.

The Directions for Task Administration section within each PAAP Task Description includes additional supports not listed among the approved general assessment accommodations. Because of the modified nature of the PAAP and the population for whom the PAAP is intended, some flexibility in the method of presentation is necessary and appropriate. It is important to remember that the use of these support mechanisms does not affect the PAAP scoring formula. They do not change what is being measured in the task.

If a student needs supports beyond those provided through approved accommodations or the flexibility that is part of every PAAP Task Description, the opportunity to use individualized Levels of Assistance is provided. Supports classified as Levels of Assistance are teacher-developed support mechanisms that, while not modifying the content being measured, assist a student in completing the task or retrieving the answer to a particular question without actually providing that answer to the student.

Levels of Assistance are determined on a three-point scale of 1–3, with each point affecting the overall score of a PAAP task. Note that as the teacher support decreases, the point score increases. These point values do not affect the student's preliminary score for the task (the percent correct). Rather, the points awarded for Levels of Assistance make up one part of the final scoring matrix, along with Level of Accuracy and LoC. The following are the Levels of Assistance by score point:

- Level of Assistance Score of 1

    - Modeling
    - Demonstrating a response similar to that desired (e.g., Teacher says, “When I turn on the faucet water comes out. This water is liquid. Actual Test Question – Look at picture. Is the water shown in the picture a solid or liquid?)”
- Level of Assistance Score of 2
    - Use of Option 2 (provided at LoC 1 when appropriate) to use fewer of the item sets multiple times in order to match the student’s knowledge
    - Limiting a student’s response (except at LoC 1) by removing one response option (e.g., multiple-choice items/problems) and reducing the response options from 3 to 2
    - Asking clarifying questions to stimulate student thought without providing clues to specific answers (e.g., “Which happened first? Show me on your board.”)
- Level of Assistance Score of 3
    - Independent
    - Providing encouragement
    - Completing task by using augmentative/alternative means of communication
    - Repeating directions
    - Reacting to student
    - Rereading a passage
    - Reminding a student to stay focused

A special field is provided on each Task Summary page where detailed information regarding the Level of Assistance for that particular task is recorded (see Figure 5-5). The teacher administering the task must first check the appropriate box indicating the Level of Assistance needed by the student (1–3). Once a box has been marked, details regarding how the assistance was given must be circled on the provided list.

Figure 5-5. 2018–19 PAAP: Level of Assistance

Level of Assistance		
Level of Assistance 1 ☐ Circle the type of assistance from the list below.	***Level of Assistance 2*** ☐ Circle the type of assistance from the list below.	***Level of Assistance 3*** ☐ Circle the type of assistance from the list below.
• Modeling • Demonstrating a response similar to that desired • Other: _____ _____	• Use of Option 2 • Limiting student's response by removing one option • Asking clarifying questions • Prompting • Cueing • Other: _____ _____	• Independent • Encouragement • Use of augmentative/alternative communication • Repeating directions • Reacting to student • Rereading passage • Reminding student to focus • Other: _____ _____

It is vital that information regarding the Level of Assistance be recorded on each Task Summary page by the teacher administering the task, as this information is essential to the scoring of the PAAP. If such information is not provided, the task may be judged as Unscorable.

Levels of Assistance not permissible are the use of "hand-over-hand" (where the teacher prompts the student by placing his or her hand over the student's hand) or any alterations to the task. Altering a task jeopardizes the integrity of the task and its alignment to the AGLEs.

In 2018–19, the Task Summary pages were available to fill in online from the Task Bank. Teachers entered the information in the online forms, printed them, and then submitted the paper copy of the Task Summary forms with the appropriate Task Description page and student work in the student's paper portfolio.

5.5 TEST SECURITY

Maintaining test security is critical to the success of the PAAP. School personnel were informed that any concerns about breaches in test security were to be reported to the school's test coordinator and/or principal immediately. The test coordinator and/or principal was responsible for immediately reporting the concern to the district superintendent and the state alternate assessment director at the Maine DOE. Test security was also strongly emphasized at the test administration workshops. The Test Administration Security Agreements are found on the Maine DOE website at: https://www.maine.gov/doe/Testing_Accountability/MECAS/materials/paap.

Additionally, principals were required to sign and submit a Principal Certification of Proper Administration (PCPA). Principals were requested to log on to a secure website to access the PCPA form. Principals were instructed to submit the form by entering a unique password, which acted as their digital signature. By submitting the form, the principal certified that the tests were administered according to the test administration procedures outlined in the *Maine PAAP Administration Handbook* and *Maine PAAP Operational Procedures 2018-19 Administration*, that each student with a disability participated in the PAAP as a result of a decision by his or her IEP team, and this decision is documented in the students' IEP plan. The principals certified that the entries submitted in each portfolio reflect the performance of the student for whom the portfolio is being submitted, and no tasks have been altered or misrepresented, that the principal has observed each teacher administering at least one task to a student, and all portfolios are being returned to Advance Education/Measured Progress (now Cognia)as prescribed on or before the scheduled deadline. These agreements are kept on file with the principal/test coordinator. PCPAs are secure material with access attainable through communication with Advance Education/Measured Progress (now Cognia).

CHAPTER 6 TEST DEVELOPMENT

The PAAP is intended to provide students with significant cognitive disabilities the opportunity to participate in a statewide assessment that is both meaningful and academically challenging. Given the wide diversity of this student population, great emphasis is placed on ensuring that the PAAP is appropriate and accessible to all students. The assessment design allows students to progress through eight Levels of Complexity (LoCs). LoC 1, the lowest LoC, represents the lowest level of knowledge and entry-level skills and therefore provides students with the most access while still maintaining an academic foundation aligned to grade-level content.

6.1 GENERAL PHILOSOPHY

The development of science for the PAAP began with face-to-face meetings over the course of three days (September 1–3, 2009) at Measured Progress (now Cognia). Each meeting consisted of Maine DOE staff, Measured Progress (now Cognia) special education project management staff, and Measured Progress (now Cognia) Content, Design & Development staff. The purpose of the meetings was to collaborate on plans for the development of tasks to finish populating the PAAP Task Bank in science.

The notes from the above-mentioned planning meetings were frequently referenced by the Content, Design & Development staff, special education specialist, and Maine DOE staff as items were drafted and reviewed. In addition to the Measured Progress (now Cognia) review process, staff from the Maine DOE and small groups of stakeholders evaluated all tasks through a task tryout process. This multistage development and review process provided ample opportunities to evaluate items for their alignment, accessibility, appropriateness, and adherence to the principles of universal design. In this way, accessibility emerged as a primary area of consideration throughout the item development process. This was critical in developing an assessment that allows for the widest range of student participation as educators seek to provide access to the general education curriculum and to foster high expectations for students with significant cognitive disabilities.

Table 6-1 indicates the full development of science tasks by LoC for the Maine PAAP Task Bank. This was completed in 2010–11.

Table 6-1. 2010–2011 PAAP: Task Development—Science

LoCs	*AGLE[1]/Indicator*	*Number of Tasks*	*Total Tasks*
1–2	D2	2 tasks per LoC[2]	4
1–8	D3, E1, E5	2 tasks per LoC[2]	48
3–4	D2	2 tasks per LoC[2]	4
5–6	D4	2 tasks per LoC[2]	4
4	E2	2 tasks per LoC[2]	2
6	E3	2 tasks per LoC[2]	2
4 & 6	E4	2 tasks per LoC[2]	4
Overall Science task revision total			68

[1] AGLE = Alternate Grade Level Expectation
[2] LoCs = Levels of Complexity as described in the AGLEs

6.2 Role of Committees in Test Development

The Advisory Committee comprised teachers, education administrators, representatives from higher education, and other agencies who advised the Maine DOE on defining the parameters of the alternate assessment. The committee was asked to review the issues related to the creation of the AGLEs, PAAP blueprint, PAAP tasks, and the achievement-level descriptors for students who are unable to participate in statewide assessments even with accommodations. Members responded to written samples and recommendations from internal groups at Measured Progress (now Cognia) and the Maine DOE regarding these areas to ensure accountability for students taking the PAAP. They also worked with the Maine DOE to determine the structures that serve as the basis for today's PAAP.

CHAPTER 7 ALIGNMENT

7.1 DESCRIPTION OF LINKAGES TO DIFFERENT CONTENT AREAS ACROSS GRADES

The Maine DOE contracted two external alignment studies to be completed by Amy S. Burkam, Lothlorien Consulting, LLC: one in March 2010 and one in June 2012. The March 2010 study was conducted in two phases. The results of the first study are documented in the 2010–11 technical report. The 2012 study results were discussed in the 2011–12 technical report.

CHAPTER 8 PAAP ADMINISTRATION TRAINING

In December 2018, the Maine DOE, in collaboration with program management at Measured Progress (now Cognia), trained teachers from across the state via a half-day in-person training in the process of constructing a PAAP. This training provided test administrators with the steps for administration of the PAAP process (see Section 8.1), a thorough review of the Alternate Grade Level Expectations (AGLEs) document, and other changes in procedures that had been made since the prior year. Beginning in June 2012, teachers' scores submitted electronically on the Task Summary page in the Task Bank were used as the first score. The second score was provided by Advance Education/Measured Progress's (now Cognia)trained scorers. This use of the Task Bank was integrated into the PAAP update presentations.

Four in-person administration trainings were held December 4-7 in Presque Isle, Orono, Augusta, and Saco. Table 8-1 outlines the number of participants who attended the trainings.

Table 8-1. 2018–19 PAAP: Administration Training Attendance Count

Workshop Location	*Attendance*
Presque Isle	8
Orono	30
Augusta	66
Saco	44

The administration training provided teachers with guidance on how to submit a PAAP for scoring. The purpose was to remind teachers of the process required to electronically submit the Task Summary pages (where teachers recorded the students' scores) via the ProFile Task Bank before the administration window closed on April 30, 2019. A PowerPoint titled "2018–19 Introduction to the PAAP" and a prerecorded Webinar was also posted on the Maine DOE website:
https://www.maine.gov/doe/Testing_Accountability/MECAS/materials/paap

8.1 STEPS FOR ADMINISTRATIONS

A detailed handbook was developed by the Maine DOE in collaboration with Measured Progress (now Cognia) as a training tool to instruct teachers on how to design and implement a PAAP. Minimal changes to the handbook occurred between the 2017–18 and 2018–19 administrations. The *2018–19 PAAP Administration Handbook*, which includes the AGLEs, was available to download from the Maine DOE's website https://www.maine.gov/doe/Testing_Accountability/MECAS/materials/paap).

The administration process, clearly outlined in the *2018–19 PAAP Administration Handbook*, is broken into steps that guide the teacher from the point of determining student eligibility to the actual

submission of the PAAP. The handbook provides detailed information to the reader on what evidence to collect and how to design a PAAP appropriate for an individual student.

The main *2018–19 PAAP Administration Handbook* sections are as follows:

- Vital Information At-a-Glance
- Introduction
- Determining the Appropriate Avenue for Student Participation in State Assessments
- Alternate Grade Level Expectations (AGLEs)
- The Task Bank
- Types of Supports
- Administering a PAAP
- Scoring a PAAP
- Reporting
- Code of Conduct
- Supplemental Materials

Announcements of the upcoming trainings and registration information were posted on the PAAP website and e-mailed to special education directors. Workshop registration was submitted through Measured Progress ’s (now Cognia) online registration application.

8.2 STEPS IN CONSTRUCTING THE PORTFOLIO

The steps and scoring ramifications for constructing the PAAP are outlined in the *2018–19 PAAP Administration Handbook* to assist teachers in planning, instructing, and assessing students taking a PAAP. The steps are:

A. Planning a PAAP

Step 1

Meet with the student’s individualized education program (IEP) team to determine the appropriate avenue of participation in the state assessment using the Maine Participation Guidelines. The Participation Guidelines, a flow chart and a participation checklist were available in the handbook to guide the IEP team to determine the most appropriate avenue of assessment to determine participation in the PAAP.

- Scoring Ramifications: Participation in the PAAP by a student who does not meet the defined guidelines will result in the student being counted as a nonparticipant in the MEA Science.

Step 2

Using the grade-level blueprint, choose the required number of AGLE/Indicators for which the student will submit a PAAP. The AGLE/Indicators will be the target of instruction for the individual student. Related instruction and assessment should be integrated with the student’s IEP.

- Scoring Ramifications: If student work is submitted for fewer than the required number of AGLE Entries, the raw score for the PAAP will be lower and may not accurately reflect the student's level of knowledge and skills. AGLE Entries submitted beyond the number required will not be scored.

Step 3

For each AGLE/Indicator required, use the PAAP AGLEs to identify the Level of Complexity (LoC) descriptors that are appropriate for inclusion in the student's instructional program. Consider the student's current level of performance and the possibilities for increasing that level through instruction as you read the PAAP LoC descriptors. The LoC should challenge the student and allow the opportunity for the student to demonstrate proficiency.

B. Registering a Student for PAAP

Step 4

Create a user account within the PAAP Task Bank. This can be done by using the registration button on the top of the Task Bank homepage. The Task Bank can be accessed by going to https://www.maine.gov/doe/Testing_Accountability/MECAS/materials/paap and clicking on the "Task Bank" button. More detailed instructions on creating your account can be found in the *Task Bank Manual* located on the homepage of the Task Bank.

Step 5

Add students to your list by entering the State Student ID (SSID) and then verifying the student name and grade upon pressing the "OK" button.

Step 6

Verify that the student information is accurate. Then use the "Add to Student List" button to register the student.

If the student information is not accurate, contact the person responsible for entering and uploading student ID data to the state student enrollment site Synergy (this may be your building secretary or other designee). If this information needs to be updated in the Task Bank or the student record is not found in the Task Bank, contact the Alternate Assessment Coordinator at the Maine DOE to make changes in the Task Bank.

C. Implementing a PAAP

Step 7

Using tasks from the Task Bank, collect student work for the required AGLE/Indicators throughout the testing window. Students may have been assessed on a task multiple times during the testing window. Submit only the required number of completed tasks for an Entry.

When the teacher records the answer on the student work template, the teacher must indicate the student response (e.g., "student pointed" on the answer line is not sufficient; you must write "student pointed to the cup").

- Scoring Ramifications: Fewer than the required number of tasks submitted for an AGLE Entry will result in the task being "Unscorable." Extra student work submitted will not be scored and may result in scorer confusion and negatively affect the scoring process for the PAAP. If there is no student response listed, the task may be "Unscorable."

Step 8

Fill out a single Entry Slip for each AGLE Entry that you are assessing for the PAAP.

- Submit three AGLE Entries in science.
- Scoring Ramifications: Student work submitted without an Entry Slip may result in scorer confusion and negatively affect the scoring process for the PAAP.

Step 9

On the Work Template, make sure information has been filled in for all sections, including the Student Response column.

- Scoring Ramifications: Work Templates that are not completely filled in may result in an inability to score the work for the Task, or even the entire AGLE Entry.

Step 10

All student work must be corrected item-by-item on the Work Template. Use an "X" for an incorrect response and a "C" for a correct response. If the student self-corrects (i.e., without any prompting, changes error), please clearly indicate this and score the student's final answer choice. Transfer the student's correct/incorrect scores to the online Task Summary page.

- Scoring Ramifications: Work that has not been corrected item-by-item will be considered "Unscorable."

Step 11

Using Levels of Assistance information, determine the Level of Assistance score that best represents the Level of Assistance earned. The teacher is required to indicate how assistance was given by checking an entry from the populated list or by writing a brief description in the "Other" section.

- Scoring Ramifications: The description is used to verify the score for this variable. Simply checking one of the boxes on the Task Summary page does not provide the scorer with sufficient information and will result in the task being "Unscorable."

Step 12

Electronically complete and submit all Task Summary pages. Information within the Level of Accuracy box and the Level of Assistance section must be populated. Refer to Levels of Assistance to determine the score.

Task Summary pages must be filled in electronically and submitted online (by April 30) using the Task Bank and be included in the portfolio. The electronic submission will result in the student's first score of the portfolio, while the paper version will assist the second scorer.

- Scoring Ramifications: Task Summary pages that are not filled in electronically and submitted online by April 30 using the Task Bank will result in the inability to score the work for the AGLE Entry.

D. Organizing a PAAP

Step 13

Assemble each AGLE Entry by attaching the required number of Task Descriptions with accompanying student work and Task Summary pages. Do not attach the following:

- More than the required number of Task Descriptions
- More than the required amount of student work and Task Summary pages
- Description cards and/or cutout graphics used for the tasks (If you would like to save these items, place them in a separate section at the end of the PAAP.)
- Scoring Ramifications: Student work submitted without an Entry Slip and/or without the required number of Task Descriptions may result in scorer confusion and negatively affect the scoring process for the PAAP. Student work submitted without the required number of Work Templates and/or the required number of Task Summary pages will result in the entry being "Unscorable." Extra Task Descriptions and/or student work submitted will not be scored and may result in scorer confusion and negatively affect the scoring process for the student's PAAP.

Step 14

Arrange each AGLE Entry in alphabetical order by AGLE and then in numerical order by Indicator. Refer to the grade-level blueprint for more details.

- Scoring Ramifications: Lack of organization may result in scorer confusion and negatively affect the scoring process.

Step 15

Print the Table of Contents (available through the Task Bank or on the PAAP website) and check that all white sections of the Entry Slips (Name and Grade), Student Work (Name and Date), and Task Summary page (Name, Date, Level of Accuracy, and Level of Assistance) have been filled in.

- Scoring Ramifications: Incomplete documentation and lack of organization can result in an inability to score the PAAP.

E. Submitting a PAAP

Step 16

Prepare the PAAP for mailing according to the directions received from Measured Progress (now Cognia) in the return materials shipment that will be sent in April. Measured Progress (now Cognia)has arranged for a one-day UPS pickup of all PAAPs during the first week of May from every school with PAAP

students. UPS will deliver the PAAPs to Advance Education/Measured Progress (now Cognia). PAAPs will be returned to schools at the start of the new school year.

- Scoring Ramifications: Any PAAPs received later than one week from the pickup date will not be scored, and students for whom late PAAPs have been submitted will be counted as nonparticipants in the MEA Science.

IMPORTANT: Sending schools are responsible for verifying that students who are tuitioned to Special Purpose Private Schools, or who are attending out-of-district programs, are being assessed.

CHAPTER 9 SCORING

One 2018–19 three-day scoring session was held at Advance Education/Measured Progress (now Cognia)in Dover, New Hampshire. Nine professionally trained scorers and three table leaders participated in the scoring session. The Advance Education/Measured Progress (now Cognia)scorers were interviewed, hired (based on MEA/PAAP-established scorer criteria), and trained for PAAP scoring. The 12 participants scored a total of 387 PAAPs.

9.1 TABLE LEADER AND SCORER RECRUITMENT AND QUALIFICATIONS

Table leaders and scorers were handpicked by Advance Education/Measured Progress (now Cognia)staff from a pool of experienced table leaders and scorers and were required to pass a qualifying set of sample portfolio entries. Scorers and table leaders were required to sign nondisclosure agreements and agree to maintain the security of PAAP materials at all times. The scorer code of conduct, which details the importance of confidentiality and bias-free scoring, was also reviewed with the scorers.

9.2 TABLE LEADER AND SCORER TRAINING

Measured Progress (now Cognia) table leaders and scorers attended a training session at Measured Progress (now Cognia) on May 28, 2019. The first half of the training session was held specifically for table leaders. They were trained on their responsibilities as table leaders, which included the online scoring application, the flow of materials at their tables, and monitoring third reads. Scorers joined the table leaders for the second half of the training session for an in-depth review of the materials.

The training included a PowerPoint presentation showing the steps required in the scoring process, from checking the student name to entering scores in the online application developed for PAAP scoring. Staff members then conducted hands-on training in the use of the online ProFile scoring application. A sample portfolio for “Liam,” a fictitious student, contained entries and was used to illustrate the scoring process. These sample entries, including potential scoring issues, were reviewed and discussed. Next, table leaders and scorers practiced using sample sets before taking online qualifying sets. All prospective table leaders and scorers qualified by earning the required scores on these sets. Prior to any scoring, table leader guidelines were reviewed to assure consistency in their understanding of the expectations. In addition, a table leader debrief occurred at the end of each scoring day to review procedures and address issues that came up during scoring.

Personnel from Measured Progress (now Cognia) and the Maine DOE were available to answer questions that arose during both the training and the actual scoring sessions. This was essential as clarifications to any scoring irregularities/rules or Alternate Grade Level Expectations (AGLEs)/Indicators

arose, as well as some initial assistance with the online scoring application. Scorers were provided with the 2018–19 AGLEs (see Appendix A), 2019 scoring instructions, 2018–19 task scoring rubric, and 2019 scoring rules.

9.3 SCORING PROCESS

PAAP scoring was conducted using the online ProFile scoring application, which was developed for this contract. The ProFile application allowed teachers' scores and scoring staff scores to be submitted online. Teachers' scores were used for the first score of record, and the scoring staff provided the second score. Teachers were required to complete Task Summary pages electronically for their students through the ProFile Task Bank, while Measured Progress 's (now Cognia) scoring staff members submitted their scores on a similar Task Summary page in the scoring application. Each PAAP was read and scored at least once by a Measured Progress (now Cognia) scorer, with some of the PAAPs being scored a third time in a double-blind fashion. (See Section 11.4 for interrater consistency.) A PAAP was also scored a third time if scorers 1 (teacher) and 2 (scoring staff) did not have exact agreement for Level of Complexity (LoC), Level of Accuracy, or Level of Assistance on any content standard entry. The third score was the final score of record for each dimension that was discrepant. The third scores were determined by the Maine DOE, Measured Progress (now Cognia) program management personnel, and senior scoring staff.

The scoring process was explained in detail to both the table leaders and the scorers. The following steps were required of all table leaders and scorers.

Step 1. Prescreen the PAAP. Scorers were to ensure that:

- the student was not known to the scorer, and
- the PAAP was organized correctly.

Step 2. Log in to the scoring application using the assigned scorer number and password. The scorer ID was attached to the PAAP, thereby identifying scorer 2.

Step 3. Verify that the student information on the portfolio matches that in the ProFile scoring application. Scorers were instructed to verify that the portfolio demographic information provided on the Verify Demographics screen (student ID number, name, grade, and district and school names) matched the information on the portfolio. If they were the same, then the scorer continued to the next step. If there were any differences, the scorer alerted senior staff to resolve the issue.

Step 4. Verify that all the required components are present. Scorers used the ProFile Verify Entries screens to determine if the portfolio contained all the requisite pieces and if the grade requirements had been met. If an AGLE/Indicator was incorrect or any portfolio pieces were identified as missing, then the scorer would indicate the problem by assigning the associated comment code (refer to Step 6, Provide comment codes) and finalize that entry.

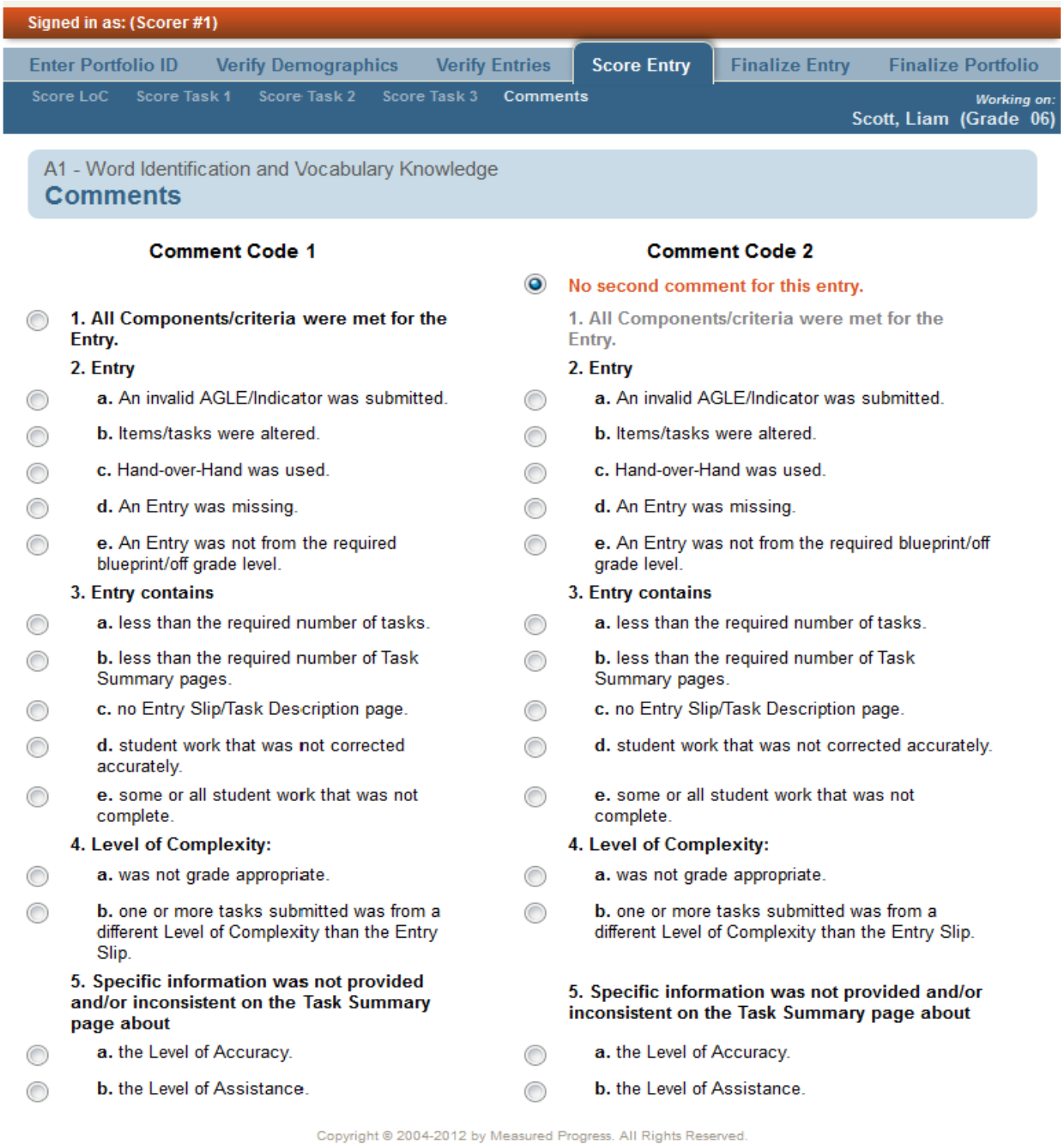
Step 5. Score each entry. If an entry was determined to be scorable, the scorer then read the individual tasks that met the requirements for an entry and scored them in ProFile on three dimensions—LoC, Level of Accuracy, and Level of Assistance.

Step 6. Provide comment codes. Scorers also received instruction on how to complete comment codes, which provide teachers with valuable feedback on the entry scores. At least one comment code must be selected in ProFile for each entry (maximum of two). Based on the totality of the entry and the information provided on the comment code sheet, the second scorer selected one or two comment codes for the entry (see Figure 9-1).

In the quality-control area, ProFile provided a real-time list of unscored and discrepant portfolios that were then located and distributed appropriately for scoring. As an added step, Measured Progress (now Cognia)personnel tracked the portfolios to ensure that all had been scored and accounted for at the end of the scoring session.

Refer to Appendix C for additional documents that were used during scoring. The 2019 PAAP Scoring Instructions describes the scoring process in greater detail than noted above. The 2019 Task Scoring Rubric provides an overview of the scores related to each dimension—LoC, Level of Accuracy, and Level of Assistance.

Figure 9-1. 2018–19 PAAP: PAAP Comment Codes



9.4 FLOW OF MATERIALS

The scoring teams used the following instructions for the flow of materials in the day-to-day scoring of the PAAPs.

Scorers:

- request a PAAP from the table leader.
- verify that the student information on the portfolio matches that in the ProFile scoring application.
- verify that all required contents of the PAAP are inside the binder.
- score according to the 2019 Scoring Instructions sheet (Appendix C).
- enter scores accurately in ProFile.
- return scored PAAP to the table leader.
- repeat this process with each PAAP.

Table Leaders:

- make sure that at least one box of unscored PAAP binders is available.
- distribute unscored PAAPs to scorers.
- perform a read-behind of each scorer's first PAAP and any scorer evaluated by a table leader as having difficulty with the process; review random PAAPs throughout the day to validate the scoring.
- meet with the scorer immediately if any problems with scoring are noticed. If problems persist, notify personnel from the Maine DOE or Measured Progress (now Cognia) .
- place the PAAP in its original envelope.
- place the envelope in the box from which it came.
- score additional PAAPs as outlined in the scoring instructions above.

9.5 SECURITY

During scoring workdays, all PAAPs remained in sight of Measured Progress (now Cognia) and Maine DOE personnel at all times. During the day, PAAPs were methodically delivered back and forth from the quality-control room to the scoring room. At the end of each day, PAAPs were stored in a locked room.

Measured Progress's (now Cognia) distribution personnel delivered the PAAPs directly to the Measured Progress (now Cognia) scoring site. After all scoring was completed, the PAAPs were returned to the Measured Progress (now Cognia) warehouse, where they were stored until September 9, 2019, when they were shipped back to their point of origin.

9.6 SCORING RUBRIC

During PAAP scoring, the 2019 PAAP task scoring rubric was used to determine the official scores of record for LoC, Level of Accuracy, and Level of Assistance.

Level of Accuracy was scored on a Likert scale of 1–4 based on the overall task percent correct score (e.g., a task percent correct score of 67% would receive an overall Level of Accuracy score of 3). Figure 9-2 demonstrates how a score of 1, 2, 3, or 4 was obtained. When scorers marked each item as correct/incorrect, ProFile automatically calculated the Level of Accuracy scores (1–4) based on the figure below.

Figure 9-2. 2018–19 PAAP: Task Score for Level of Accuracy

Task Score for Level of Accuracy			
1	2	3	4
Student work was not corrected. Student work related to the Task was completed with a score of 0 – 19%.	Student work related to the Task was completed with a score of 20 – 60%.	Student work related to the Task was completed with a score of 61 – 84%.	Student work related to the Task was completed with a score of 85 – 100%.

Level of Assistance was scored on a scale of Unscorable (receiving a score of 0) to 3, based on the approved accommodations outlined in Figure 9-3. The scorer entered the Level of Assistance and the type of support provided from the drop-down list below each Level of Assistance.

Figure 9-3. 2018–19 PAAP: Task Score for Level of Assistance

Task Score for Level of Assistance			
Unscorable	1	2	3
➢ Hand-over-hand ➢ Altering items/tasks (task no longer connects to the AGLE)	➢ Modeling ➢ Demonstrating a response similar to the desired response	➢ Use of **Option 2** (LoC 1 only) to use fewer of the item sets multiple times to match student knowledge ➢ Limiting a student's response (outside of LoC 1 at Option 2) by removing one response option ➢ Use of clarifying questions to stimulate student thought to the specific task without providing clues to specific answers	➢ Independent ➢ Providing encouragement ➢ Completing tasks by using augmentative/alternate means of communication ➢ Repeating directions ➢ Reacting to a student ➢ Rereading a passage ➢ Reminding a student to stay focused

The 2019 PAAP task scoring rubric is shown in Figure 9-4.

Figure 9-4. 2018–19 PAAP: Task Scoring Rubric

MAINE PAAP Personalized Alternate Assessment Portfolio

TASK SCORING RUBRIC

Task Score for Level of Complexity								
Unscorable	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
The PAAP Task did not meet all requirements. **Reasons for Unscorables:** ✓ Level of Complexity does not match Entry Slip and is not consistent with other tasks. ✓ Level of Complexity is not grade appropriate.								

Task Score for Level of Accuracy			
1	**2**	**3**	**4**
Student work related to the Task was completed with a score of 0 - 19%.	Student work related to the Task was completed with a score of 20 - 60%.	Student work related to the Task was completed with a score of 61 - 84%.	Student work related to the Task was completed with a score of 85 - 100%.

Task Score for Level of Assistance			
Unscorable	**1**	**2**	**3**
Hand-over-hand Altering items/tasks beyond removing a choice (task no longer connects to the AGLE)	Modeling Demonstrating a response similar to the desired response Other	Use of **Option 2** (LoC 1 only) to use fewer of the item sets multiple times to match student knowledge Limiting a student's response (outside of LoC 1 at Option 2) by removing one response option Use of clarifying questions to stimulate student thought to the specific task without providing clues to specific answers Other	Independent Providing encouragement Completing tasks by using augmentative/alternate means of communication Repeating directions Reacting to a student Rereading a passage Reminding a student to stay focused Other

09262012

After each PAAP was scored, a table leader from Measured Progress (now Cognia) removed the PAAP from its envelope to confirm that it corresponded with the student identified on the envelope. The PAAP was then inserted in its envelope and returned to the box.

Then the box of PAAPs was returned to the quality-control room where it remained unless a PAAP was identified via the ProFile scoring application as needing a third score. At this time, the PAAP was provided to either a Measured Progress (now Cognia) program manager or a staff member of the Maine DOE for a third read.

When the person doing the third read entered the PAAP identification number in ProFile for a third score, the application displayed the scoring dimension(s) in disagreement on the screen. The score resulting from the third read became the score of record. At this point, the PAAP was considered complete and filed in its box.

9.7 Calculation of Reported Scores

After the scoring process was completed, students' scores on each entry were calculated based on a formula that combines their LoC, Level of Accuracy, and Level of Assistance scores for each of the tasks in that entry. The formula weights the LoC score more heavily than the other two dimension scores. The overall score of record is then the sum of the entry scores. Because of the use of the formula, there may be multiple ways that a student can attain a given total score. Complete details of how reported raw scores are calculated are provided in Appendix D.

CHAPTER 10 CLASSICAL ITEM ANALYSIS

As noted in Brown (1983), "A test is only as good as the items it contains." A complete evaluation of a test's quality must include an evaluation of each item. Both *Standards for Educational and Psychological Testing* (AERA et al., 2014) and *Code of Fair Testing Practices in Education* (Joint Committee on Testing Practices, 2004) include standards for identifying quality items. While the specific statistical criteria identified in these publications were developed primarily for general—not alternate—assessment, the principles and some of the techniques apply within the alternate assessment framework as well.

Both qualitative and quantitative analyses were conducted to ensure that Maine PAAP items met these standards. Qualitative analyses are described in earlier sections of this report; this section focuses on the quantitative evaluations. The statistical evaluations discussed are difficulty indices and discrimination (item-test correlations), structural relationships (correlations among the dimensions), and bias and fairness. The item analyses presented here are based on the statewide administration of the 2018–19 PAAP.

10.1 DIFFICULTY AND DISCRIMINATION

For the purpose of calculating item statistics, the two dimension scores on each task (Level of Accuracy and Level of Assistance) can be considered similar to those for traditional test items. *Difficulty* was defined as the average proportion of points achieved on an item and was measured by obtaining the average score on an item and dividing by the maximum score for the item. Using this definition, all items were evaluated in terms of item difficulty according to standard classical test theory practices. PAAP tasks are scored polytomously, such that a student can achieve a score of 1, 2, 3, or 4 for Level of Accuracy and a score of 1, 2, or 3 for Level of Assistance. By computing the difficulty index as the average proportion of points achieved, the items are placed on a scale that ranges from 0.0 to 1.0. Although the *p*-value is traditionally described as a measure of difficulty (as it is described here), it is properly interpreted as an easiness index, because larger values indicate easier items.

An index of 0.0 indicates that all students received no credit for the item, and an index of 1.0 indicates that all students received full credit for the item. Items that have either a very high or very low difficulty index are considered potentially problematic, because they are either so difficult that few students get them right or so easy that nearly all students get them right. In either case, such items should be reviewed for appropriateness for inclusion on the assessment. If an assessment were composed entirely of very easy or very hard items, all students would receive nearly the same scores, and the assessment would not be able to differentiate high-ability students from low-ability students.

It is worth mentioning that using a norm-referenced criterion such as *p*-values to evaluate test items is somewhat contradictory to the purpose of a criterion-referenced assessment like the PAAP. Criterion-referenced assessments are primarily intended to provide evidence on student progress relative to a standard

rather than to differentiate among students. Thus, the generally accepted criteria regarding classical item statistics are only cautiously applicable to the PAAP.

A desirable feature of an item is that higher-ability students perform better on the item than do lower-ability students. The correlation between student performance on a single item and total test score is a commonly used measure of this characteristic of an item. Within classical test theory, this item-test correlation is referred to as the item's "discrimination," because it indicates the extent to which successful performance on an item discriminates between high and low scores on the test. The discrimination index used to evaluate PAAP items was the Pearson product-moment correlation. The theoretical range of this statistic is -1.0–1.0.

Discrimination indices can be thought of as measures of how closely an item assesses the same knowledge and skills assessed by other items contributing to the criterion total score. That is, the discrimination index can be thought of as a measure of construct consistency. In light of this interpretation, the selection of an appropriate criterion total score is crucial to the interpretation of the discrimination index. For the PAAP, the test total score was used as the criterion score.

A summary of the item difficulty and item discrimination statistics for each content area/grade combination is presented in Table 10-1. The mean difficulty values shown in the table indicate that, overall, students performed well on the items on the PAAP. In contrast to alternate assessments, the difficulty values for assessments designed for the general population tend to be in the 0.4–0.7 range for the majority of items. Because the nature of alternate assessments is different from that of general assessments, and because very few guidelines exist as to criteria for interpreting these values for alternate assessments, the values presented in Table 10-1 should not be interpreted to mean that the students performed better on the PAAP than the students who took general assessments. An additional factor, as mentioned above, is that item statistics are calculated from students' Level of Accuracy and Level of Assistance scores. Students' overall scores, on the other hand, are based on the Level of Accuracy and Level of Assistance scores along with the Level of Complexity (LoC). A formula that combines the three dimensions, weighting LoC more heavily, is used to compute the students' score of record. Looking at the *p*-values in isolation would suggest that students' reported scores would all be very high; however, use of the formula results in reported scores that show greater spread across the range of obtainable scores than would be expected based on the *p*-values alone (see Appendix D for complete details on how the score of record is calculated).

Also shown in Table 10-1 are the mean discrimination values. Because the majority of students received high scores on the tasks, the discrimination indices are somewhat lower than one might expect. This is an artifact of how discrimination is calculated: If all the students get an item correct, there is little variability in the criterion scores that can be differentiated. As with the item difficulty values, because the nature and use of the PAAP are different from those of a general assessment, and because very few guidelines exist as to criteria for interpreting these values for alternate assessments, the statistics presented in Table 10-1 should be interpreted with caution.

Table 10-1. 2018–19 PAAP: Summary of Item Difficulty and Discrimination Statistics by Content Area and Grade

Content Area	*Grade*	*Number of Items*	p-*Value*		*Discrimination*	
			Mean	*Standard Deviation*	*Mean*	*Standard Deviation*
Science	5	48	0.77	0.11	0.54	0.19
	8	72	0.74	0.16	0.50	0.31
	HS	96	0.77	0.10	0.52	0.33

In addition to the item difficulty and discrimination summaries presented above, item-level classical statistics and item-level score distributions were also calculated. Item-level classical statistics are provided in Appendix E; item difficulty and discrimination values are presented for each item. Item-level score distributions (i.e., the percentage of students who received each score point) are provided in Appendix F for each item.

10.2 STRUCTURAL RELATIONSHIP

By design, the achievement-level classification of the PAAP is based on two of the three dimensions (accuracy and assistance). The third dimension (complexity) is used at the time of administering the assessment to determine the specific sets of tasks appropriate for the student. As with any assessment, it is important that these dimensions be carefully examined. This was achieved by exploring the relationships among student dimension scores with Pearson correlation coefficients. A very low correlation (near 0) would indicate that the dimensions are not related; a low negative correlation (approaching -1.00), that they are inversely related (i.e., that a student with a high score on one dimension had a low score on the other); and a high positive correlation (approaching 1.00), that the information provided by one dimension is similar to that provided by the other dimension.

The average correlations between Level of Accuracy and Level of Assistance by content area and grade are shown in Table 10-2.

Table 10-2. 2018–19 PAAP: Average Correlations Between Level of Accuracy and Level of Assistance by Content Area and Grade

Content Area	*Grade*	*Number of Items*	*Average Correlation*	*Correlation Standard Deviation*
Science	5	24	0.44	0.31
	8	36	0.41	0.42
	HS	42	0.41	0.40

10.3 BIAS/FAIRNESS

Code of Fair Testing Practices in Education (Joint Committee on Testing Practices, 2004) explicitly states that subgroup differences in performance should be examined when sample sizes permit, and actions should be taken to make certain that differences in performance are due to construct-relevant rather than construct-irrelevant factors. *Standards for Educational and Psychological Testing* (AERA et al., 2014) includes similar guidelines.

When appropriate, the standardization differential item functioning (DIF) procedure (Dorans & Kulick, 1986) is used to identify items for which subgroups of interest perform differently, beyond the impact of differences in overall achievement. However, because of the small number of students who take the PAAP, and because those students take different combinations of tasks, it was not possible to conduct DIF analyses. This is because conducting DIF analyses using groups of fewer than 200 students would result in inflated type I error rates.

Although it is not possible to run quantitative analyses of item bias for PAAP, fairness is addressed through Cognia's standard item development and review procedures, described in detail earlier in this report. These procedures, which are modeled on the recommendations laid out in *Standards for Educational and Psychological Testing* (AERA et al., 2014), are designed to ensure that the test is free of any insensitive or offensive material. All tasks that are available to teachers in the standardized Task Bank have been through these comprehensive bias and sensitivity reviews.

Issues of fairness are also addressed in the PAAP scoring procedures. Chapter 9 of this report describes in detail the scoring rubrics used, selection and training of scorers, and scoring quality-control procedures. These processes ensure that bias due to differences in how individual scorers award scores is minimized.

CHAPTER 11 CHARACTERIZING ERRORS ASSOCIATED WITH TEST SCORES

The main use of the PAAP scores is for school-, district-, and state-level accountability in the federal Every Student Succeeds Act (ESSA) and state accountability systems. The students are classified as Well Below State Expectations, Below State Expectations, At State Expectations, and Above State Expectations, and they are included in the state's accountability calculation.

As with the classical item statistics presented in the previous chapter, the two dimension scores on each task (Level of Accuracy and Level of Assistance) were used as the item scores for purposes of calculating reliability estimates.

11.1 RELIABILITY

In the previous chapter, individual item characteristics of the 2018–19 Maine PAAP were presented. Although individual item performance is an important focus for evaluation, a complete evaluation of an assessment must also address the way in which items function together and complement one another. Any measurement includes some amount of measurement error. No academic assessment can measure student performance with perfect accuracy; some students will receive scores that underestimate their true ability, and other students will receive scores that overestimate their true ability. Items that function well together produce assessments that have less measurement error (i.e., the error is small on average). Such assessments are described as "reliable."

There are a number of ways to estimate an assessment's reliability. One approach is to split all test items into two groups and then correlate students' scores on the two half-tests. This is known as a split-half estimate of reliability. If the two half-test scores correlate highly, the items on them are likely measuring very similar knowledge or skills. It suggests that measurement error will be minimal.

The split-half method requires psychometricians to select items that contribute to each half-test score. This decision may have an impact on the resulting correlation, since each different possible split of the test halves will result in a different correlation. Another problem with the split-half method of calculating reliability is that it underestimates reliability, because test length is cut in half. All else being equal, a shorter test is less reliable than a longer test. Cronbach (1951) provided a statistic, alpha (α), that avoids the shortcomings of the split-half method by comparing individual item variances to total test variance.

Cronbach's α was used to assess the reliability of the 2018–19 Maine PAAP tests. The formula is as follows:

$$\alpha \equiv \frac{n}{n-1}\left[1 - \frac{\sum_{i=1}^{n} \sigma^2_{(Y_i)}}{\sigma^2_x}\right],$$

where
i indexes the item,
n is the number of items,
$\sigma^2_{(Y_i)}$ represents individual item variance, and
σ^2_x represents the total test variance.

Table 11-1 presents raw score descriptive statistics (maximum possible score, average, and standard deviation), Cronbach's α coefficient, and raw score standard error of measurement (SEM) for each content area and grade. The reliability of a test can also be exhibited in terms of the SEMs, which can facilitate the interpretation of individual scores. With any given observed raw score point, the reasonable limits of the true score for the examinees can be calculated by using the SEMs. For a more detailed description about the use of SEMs, the reader is referred to Gulliksen (1950) or Anastasi and Urbina (1997). SEM was also used to assess the reliability of the 2018–19 Maine PAAP tests:

$$SEM \equiv \sigma^2_x \sqrt{1-\alpha},$$

where
σ^2_x represents the total test variance, and
α represents the reliability coefficient, Cronbach's alpha.

Table 11-1. 2018–19 PAAP: Raw Score Descriptive Statistics, Cronbach's Alpha, and SEM by Content Area and Grade

Content Area	*Grade*	*Number of Students*	*Raw Score*			*Alpha*	*SEM*
			Maximum	*Mean*	*Standard Deviation*		
Science	5	135	69	40.76	16.13	0.86	6.04
	8	131	99	57.60	26.88	0.89	8.92
	HS	120	129	70.08	34.77	0.89	11.53

An alpha coefficient toward the high end is taken to mean that the items are likely measuring very similar knowledge or skills (i.e., they complement one another and suggest a reliable assessment).

11.2 Subgroup Reliability

The reliability coefficients discussed in the previous section were based on the overall population of students who took the 2018–19 PAAP. Subgroup Cronbach's α's were calculated using the formula defined in

the section above with only the members of the subgroup in question included in the computations. These statistics are reported in Appendix G. Note that statistics are only reported for subgroups with at least 10 students.

For several reasons, the results of this section should be interpreted with caution. First, inherent differences between grades and content area preclude making valid inferences about the quality of a test based on statistical comparisons with other tests. Second, reliabilities are dependent not only on the measurement properties of a test but also on the statistical distribution of the studied subgroup. For example, it can be readily seen in Appendix G that subgroup sample sizes may vary considerably, which results in natural variation in reliability coefficients. Or α, which is a type of correlation coefficient, may be artificially depressed for subgroups with little variability (Draper & Smith, 1998). Third, there is no industry standard to interpret the strength of a reliability coefficient, and this is particularly true when the population of interest is a single subgroup.

11.3 Decision Accuracy and Consistency

While related to reliability, the accuracy and consistency of classifying students into performance categories are even more important statistics in a standards-based reporting framework (Livingston & Lewis, 1995). Unlike generalizability coefficients, decision accuracy and consistency (DAC) can usually be computed with the data currently available for most alternate assessments. For every 2018–19 PAAP, each student was classified into one of the following achievement levels: Well Below State Expectations, Below State Expectations, At State Expectations, and Above State Expectations. However, because of the small testing population for the PAAP, it was not possible to calculate DAC based on classification into the four achievement levels; instead, the categories were collapsed into Proficient or Not Proficient. Because the Proficient cut is what is used for state and federal accountability purposes, results of DAC are most critical for these two categories. This section of the report explains the methodologies used to assess the reliability of classification decisions and presents the results.

Accuracy refers to the extent to which decisions based on test scores match decisions that would have been made if the scores did not contain any measurement error. Accuracy must be estimated, because errorless test scores do not exist. Consistency measures the extent to which classification decisions based on test scores match the decisions based on scores from a second, parallel form of the same test. Consistency can be evaluated directly from actual responses to test items if two complete and parallel forms of the test are given to the same group of students. In operational test programs, however, such a design is usually impractical. Instead, techniques have been developed to estimate both the accuracy and consistency of classification decisions based on a single administration of a test. The Livingston and Lewis (1995) technique was used for the 2018–19 PAAP because it is easily adaptable to all types of testing formats, including mixed-format tests.

The accuracy and consistency estimates reported in Appendix H make use of "true scores" in the classical test theory sense. A true score is the score that would be obtained if a test had no measurement error. Of course, true scores cannot be observed and so must be estimated. In the Livingston and Lewis (1995) method, estimated true scores are used to categorize students into their "true" classifications.

For the 2018–19 PAAP, after various technical adjustments (described in Livingston & Lewis, 1995), a two-by-two contingency table of accuracy was created for each grade, where cell [i, j] represented the estimated proportion of students whose true score fell into classification i (where i = 1 or 2) and observed score into classification j (where j = 1 or 2). The sum of the diagonal entries (i.e., the proportion of students whose true and observed classifications matched) signified overall accuracy.

To calculate consistency, true scores were used to estimate the joint distribution of classifications on two independent, parallel test forms. Following statistical adjustments per Livingston and Lewis (1995), a new two-by-two contingency table was created for each grade and populated by the proportion of students who would be categorized into each combination of classifications according to the two (hypothetical) parallel test forms. Cell [i, j] of this table represented the estimated proportion of students whose observed score on the first form would fall into classification i (where i = 1 or 2) and whose observed score on the second form would fall into classification j (where j = 1 or 2). The sum of the diagonal entries (i.e., the proportion of students categorized by the two forms into exactly the same classification) signified overall consistency.

Another way to measure consistency is to use Cohen's (1960) coefficient κ (kappa), which assesses the proportion of consistent classifications after removing the proportion of consistent classifications that would be expected by chance. It is calculated using the following formula:

$$\kappa = \frac{(\text{Observed agreement}) - (\text{Chance agreement})}{1 - (\text{Chance agreement})} = \frac{\sum_i C_{ii} - \sum_i C_{i.} C_{.i}}{1 - \sum_i C_{i.} C_{.i}},$$

where

$C_{i.}$ is the proportion of students whose observed achievement level would be Level i (where i = 1 or 2) on the first hypothetical parallel form of the test;

$C_{.i}$ is the proportion of students whose observed achievement level would be Level i (where i = 1 or 2) on the second hypothetical parallel form of the test; and

C_{ii} is the proportion of students whose observed achievement level would be Level i (where i = 1 or 2) on both hypothetical parallel forms of the test.

Because κ is corrected for chance, its values are lower than are other consistency estimates.

The accuracy and consistency analyses described here are provided in Table H-1 of Appendix H. The table includes overall accuracy and consistency indices, including kappa. Accuracy and consistency values conditional on achievement level are also given. For these calculations, the denominator is the proportion of students associated with a given achievement level. For example, the conditional accuracy value is 0.81 for Not Proficient for science grade 5. This figure indicates that among the students whose true scores placed them in this classification, 81% would be expected to be in this classification when categorized according to

their observed scores. Similarly, a consistency value of 0.74 indicates that 74% of students with observed scores in the Not Proficient category would be expected to score in this classification again if a second, parallel test form was used. Table H-2 in Appendix H provides overall accuracy, consistency, false positive rate, and false negative rate. Table H-2 also provides these statistics conditional on each cutpoint. (A false positive is the proportion of students whose observed scores were above the cut and whose true scores were below the cut. A false negative is the proportion of students whose observed scores were below the cut and whose true scores were above the cut.)

Note that, as with other methods of evaluating reliability, DAC statistics calculated based on small groups can be expected to be lower than those calculated based on larger groups. For this reason, the values presented in Appendix H should be interpreted with caution. In addition, it is important to remember that it is inappropriate to compare DAC statistics between grades.

11.4 INTERRATER CONSISTENCY

Chapter 9 of this report describes in detail the processes that were implemented to monitor the quality of the hand-scoring of student responses for polytomous items. One of these processes was double-blind scoring of all student responses. Results of the double-blind scoring were used during scoring to identify scorers who required retraining or other intervention and are presented here as evidence of the reliability of the PAAP. A summary of the interrater consistency results is presented in Table 11-2. Results in the table are collapsed across the tasks bycontent area, grade, and number of score categories (3 for Level of Assistance and 4 for Level of Accuracy). The table shows the number of included scores, the percent exact agreement, the percent adjacent agreement, the correlation between the first two sets of scores, and the percent of responses that required a third score. This same information is provided at the item level in Appendix I.

Table 11-2. 2018–19 PAAP: Summary of Interrater Consistency Statistics Collapsed Across Items by Content Area and Grade

Content Area	*Grade*	*Number of Items*	*Number of*		*Percent*		*Correlation*	*Percent of Third Scores*
			Score Categories	*Included Scores*	*Exact*	*Adjacent*		
Science	5	24	3	753	97.34	1.06	0.90	2.92
		24	4	753	88.05	7.30	0.84	12.48
	8	36	3	713	98.18	1.26	0.94	1.82
		36	4	713	84.85	12.34	0.84	15.15
	HS	48	3	670	98.96	0.60	0.97	1.34
		48	4	670	88.06	9.85	0.85	12.39

CHAPTER 12 COMPARABILITY (SCALING AND EQUATING)

12.1 COMPARABILITY OF SCORES ACROSS YEARS

In administering the PAAP, teachers select tasks from a standardized Task Bank, following the test blueprints. Use of the Task Bank and blueprints ensures that the assessment as it is administered is appropriate for the individual needs of the student being assessed and that the required Alternate Grade Level Expectations (AGLEs) are covered. The process enables teachers to customize the assessment for individual students while ensuring comparability across years through the use of the same blueprints, tasks, and scoring rubrics from year to year. Additionally, comparability is ensured through the scoring process: Scoring occurs at Measured Progress (now Cognia), using the same scoring rubrics each year in addition to Measured Progress 's (now Cognia) standard scoring quality-control processes, as described in Chapter 9. Additional processes to ensure across-year comparability include calculation of reported scores and categorization into achievement levels, as described in the following sections.

12.1.1 Reported Scores

Students' entry scores are calculated based on a formula that combines their Level of Complexity (LoC), Level of Accuracy, and Level of Assistance scores for each of the tasks in a given entry. The formula weights the LoC score more heavily than the other two dimension scores. The overall score for science is then the sum of the entry scores. Because of the use of the formula, there may be multiple ways that a student can attain a given total score. Complete details of how reported raw scores are calculated are provided in Appendix D. Use of this formula ensures that the meaning of the reported raw scores will remain constant from year to year.

Graphs of the cumulative reported score distributions for 2017, 2018, and 2019 are provided in Appendix J. Note that the graphs show the proportion of students at or below each score; thus, at any given score point, the lowest line in a given graph indicates that the proportion of students scoring above that point is greatest for the year corresponding to that line. For example, in the graph for grade 8 science (bottom of Figure J-1), at the score point of 58 (i.e., Cut 2), the curve for 2018–19 is the highest, indicating that the proportion of students scoring at or above 58 is smallest for 2017–18.

12.1.2 Standard Setting

A complete standard setting was conducted for the PAAP on June 28–30, 2010. Using a rubric-based process that was supplemented with student work samples (bodies of work), standards were set for science (grades 5, 8, and 11). Although teachers are required to follow the test blueprint, they can choose which tasks

to use from the centralized Task Bank. Therefore, different students take different combinations of tasks. For this reason, a rubric-based method of standard setting was appropriate for the PAAP. Details of the standard setting procedures can be found in the standard setting report, which is posted on the Maine DOE website: https://www.maine.gov/doe/Testing_Accountability/MECAS/results/reports. To ensure continuity of score reporting across years, the cuts that were established at the standard setting meeting will continue to be used in future years, until it is necessary to reset standards. The raw score cutpoints for the PAAP as established via standard setting are presented in Table 12-1.

Table 12-1. 2018–19 PAAP: Cut Scores on the Raw Score Metric by Content Area and Grade

Content Area	*Grade*	*Raw Score*			*Raw Score*	
		Cut 1	*Cut 2*	*Cut 3*	*Minimum*	*Maximum*
Science	5	24	45	66	0	69
	8	33	58	93	0	99
	HS	50	87	127	0	129

12.2 Linkages Across Grades

In developing the PAAP, a content-based approach for addressing continuity across grades was implemented. Specifically, issues of continuity were addressed in the following processes: (1) development, (2) administration, and (3) standard setting.

As described in Chapter 5, the AGLEs describe the content to be included in students' instructional programs for each grade level. The AGLEs are based on the standards/grade-level expectations assessed by Maine's general assessments (MEA cience) but have been reduced in depth and breadth. The AGLEs are designed to follow a continuum of skills that increase across grades. The tasks, in turn, have been designed to map onto the AGLEs by measuring grade-specific content and skills. These tasks, along with blueprints, which have also been designed to reflect the continuum reflected in the AGLEs, ensure that each portfolio builds on the appropriate knowledge and skills, thereby reflecting the desired continuity across grades.

During administration, the blueprint serves as a guide to teachers when selecting tasks that are appropriate for a given student. As with other aspects of the development and administration of the PAAP, use of the blueprints and the LoCs ensures that the student is being assessed at a level that is appropriate for his or her grade level and individual needs, and that the tasks to which a student is exposed follow a continuum from year to year. Thus, linkages across grades are built into the design of the portfolio.

Finally, the continuity of the PAAP across grades was further verified through the standard setting procedures. The achievement-level descriptors used for standard setting were based on the student expectations as delineated in the AGLEs. Achievement across grades, therefore, was expected to follow the continuum established by the AGLEs and thus to reflect a higher level of cognition as the grades increased.

Following the standard setting meeting, the resulting cutpoints were critically evaluated by experts at the Maine DOE to ensure that state expectations reflected the desired increase in cognition across grades. In addition, the percentages of students scoring at or above state expectations in each grade were examined for coherence from one grade to the next.

CHAPTER 13 VALIDITY

The purpose of this report is to describe several technical aspects of the PAAP in an effort to contribute to the accumulation of validity evidence to support PAAP score interpretations. Because the combination of a test and its scores, not just the test itself, is evaluated for validity, this report presents documentation to substantiate intended interpretations (AERA et al., 2014). Each of the chapters in this report contributes important information to the validity argument by addressing one or more of the following aspects of the PAAP: test development, test administration, scoring, item analyses, reliability, achievement levels, and reporting.

The PAAP assessment is based on, and aligned to, Maine's content standards and Alternate Grade Level Expectations (AGLEs) in science. The PAAP results are intended to provide inferences about student achievement on Maine's science standards and AGLEs; these achievement inferences are meant to be useful for program and instructional improvement and as a component of school accountability.

Standards for Educational and Psychological Testing (AERA et al., 2014) provides a framework for describing sources of evidence that should be considered when constructing a validity argument. These sources include evidence based on the following five general areas: test content, response processes, internal structure, relationship to other variables, and consequences of testing. Although each of these sources may speak to a different *aspect* of validity, the sources are not distinct *types* of validity. Instead, each contributes to a body of evidence about the comprehensive validity of score interpretations.

13.1 EVIDENCE BASED ON TEST DEVELOPMENT AND STRUCTURE

A measure of test content validity is how well the assessment tasks represent the curriculum and standards for each content area and grade level. This is informed by the task development process, including how the AGLEs and test blueprints align to the curriculum and standards. Viewed through this lens provided by the content standards, evidence based on test content was extensively described in Chapters 5–7. Item alignment with Maine's content standards, AGLEs, and Levels of Complexity (LoCs), as well as review processes for item bias, sensitivity, and content appropriateness, are components of validity evidence based on test content. As discussed earlier, all PAAP tasks are aligned by Maine educators to the Maine Learning Results Science standards, AGLEs, and LoCs, and all tasks undergo several rounds of review for content fidelity and appropriateness.

Evidence based on internal structure is presented in the discussions of item analyses and reliability in Chapters 10 and 11. Technical characteristics of the internal structure of the assessments are presented in terms of classical item statistics (item difficulty, item-test correlation), correlations between the dimensions (Level of Accuracy and Level of Assistance), fairness/bias, and reliability, including alpha coefficients, interrater consistency, and decision accuracy and consistency (DAC).

13.2 Other Evidence

The training and administration information in Chapter 8 describes the steps taken to train the teachers/test administrators on procedures for constructing and administering the PAAP. Tests are constructed and administered according to state-mandated standardized procedures, as described in the *2018–19 PAAP Administration Handbook*. These efforts to provide thorough training opportunities and materials help maximize consistency among teachers, which enhances the quality of test scores and, in turn, contributes to validity.

Evidence on scoring the PAAP is provided in Chapter 9. Procedures for training and monitoring hand-scoring of the PAAP maximize scoring consistency and thus contribute to validity.

Evidence on comparability of scores, both across years and across grades, is provided in Chapter 12. Information is provided on the calculation of students' reported scores as well as the establishment of performance standards that enabled reporting of achievement-level scores. In addition, information about consistency and meaningfulness of test score information across grade levels is provided. All these processes maximize accuracy and clarity of score information that is provided to the public and, in this way, enhance validity.

13.3 Future Directions

To further support the validity argument, additional studies to provide evidence regarding the relationship of PAAP results to other variables might include an analysis of the extent to which scores from the PAAP assessments converge with other measures of similar constructs and the extent to which they diverge from measures of different constructs. Relationships among measures of the same or similar constructs can sharpen the meaning of scores and appropriate interpretations by refining the definition of the construct.

The evidence presented in this manual supports inferences related to student achievement on the content represented on the AGLEs for science for the purposes of program and instructional improvement, and as a component of school accountability.

REFERENCES

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anastasi, A., & Urbina, S. (1997). *Psychological Testing* (7th ed.). Upper Saddle River, NJ: Prentice-Hall.

Brown, F. G. (1983). *Principles of educational and psychological testing* (3rd ed.). Fort Worth, TX: Holt, Rinehart and Winston.

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

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

Dorans, N. J., & Kulick, E. (1986). Demonstrating the utility of the standardization approach to assessing unexpected differential item performance on the Scholastic Aptitude Test. *Journal of Educational Measurement*, *23*, 355–368.

Draper, N. R., & Smith, H. (1998). *Applied regression analysis* (3rd ed.). New York, NY: John Wiley and Sons.

Gulliksen, H. (1950). *Theory of Mental Tests*. New York, NY: John Wiley and Sons.

Joint Committee on Testing Practices. (2004). *Code of fair testing practices in education*. Washington, DC: Retrieved from www.apa.org/science/programs/testing/fair-code.aspx..

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

APPENDICES

APPENDIX A—2019 ALTERNATE GRADE LEVEL EXPECTATIONS

Alternate Grade Level Expectations

For

MEA (Alternate) - Science

PAAP

Personalized Alternate Assessment Portfolio

Based on Maine's Accountability Standards, Chapter 131

Science

Maine's 2007 Learning Results

©2014 Maine Department of Education. All Rights Reserved.

TABLE OF CONTENTS

SECTION **PAGE**

Blueprint of Required AGLE Indicators 2

Science AGLE D — The Physical Setting

D1 — Universe and Solar System
Levels of Complexity 1–8 3

D2 — Earth
Levels of Complexity 1–8 4

D3 — Matter and Energy
Levels of Complexity 1–8 5

D4 — Force and Motion
Levels of Complexity 1–8 6

Science AGLE E — The Living Environment

E1 — Biodiversity
Levels of Complexity 1–8 7

E2 — Ecosystems
Levels of Complexity 1–8 8

E3 — Cells
Levels of Complexity 1–8 9

E4 — Heredity and Reproduction
Levels of Complexity 1–8 10

E5 — Evolution
Levels of Complexity 1–8 11

The PAAP Blueprint of Required AGLE Indicators

Grade Level	Science
3	
4	
5	D1, D2, E2
6	
7	
8	D4, E3, E4
3rd Year High School	D3, E1, E5

Maine's Accountability Standards, Chapter 131
The Physical Setting – Universe and Solar System

Science AGLE/Indicator — D1

Student understands the universal nature of matter, energy, force, and motion, and identifies how these relationships are exhibited in Earth Systems, in the solar system, and throughout the universe by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
describing or otherwise demonstrating understanding of the positions or apparent motions of different objects in our solar system and what these objects look like from Earth by...			
doing the following: • identifying night and day.	**doing <u>both</u> of the following:** • identifying pictures of night and day, **AND** • identifying the Sun and Earth's Moon.	**doing the following:** • identifying the position of the Sun at different times by drawing or otherwise describing the movement of the Sun across the sky.	**doing <u>both</u> of the following:** • identifying the position of the Sun at different times by drawing or otherwise describing the movement of the Sun across the sky, **AND** • drawing or identifying different phases of the Moon.

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
See Extended Learning AGLEs	**See Extended Learning AGLEs**	**See Extended Learning AGLEs**	**See Extended Learning AGLEs**

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — D2
The Physical Setting – Earth

Student understands the universal nature of matter, energy, force, and motion, and identifies how these relationships are exhibited in Earth Systems, in the solar system, and throughout the universe by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
describing the properties of Earth's surface materials, the processes that change them, and cycles that affect Earth by...			
doing the following: • identifying sunny, rainy, snowy, and/or windy weather through observation.	**doing the following:** • matching pictures to the type of weather they depict.	**doing the following:** • identifying the different forms that water can take in the weather.	**doing <u>one</u> of the following:** • matching weather to the effects it can have on the surface of Earth (erosion or weathering), and/or • identifying factors that can influence temperature in the environment (day/night cycle, cloud cover, and presence of a star).

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
See Extended Learning AGLEs	**See Extended Learning AGLEs**	**See Extended Learning AGLEs**	**See Extended Learning AGLEs**

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — D3
The Physical Setting – Matter and Energy

Student understands the universal nature of matter, energy, force, and motion, and identifies how these relationships are exhibited in Earth Systems, in the solar system, and throughout the universe by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
describing properties of objects and materials before and after they undergo a change or interaction by...			
doing the following: • matching objects based on one physical property.	**doing the following:** • identifying which object in a group has a specific physical property.	**doing the following:** • sorting objects into groups using one or more physical properties.	**doing <u>both</u> of the following:** • describing the physical properties of objects and materials **AND** • using observable characteristics to describe changes in the physical properties of materials when mixed, heated, frozen, or cut.

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
describing physical and chemical properties of matter, interactions and changes in matter, and transfer of energy through matter by...		**describing the structure, behavior, and interactions of matter at the atomic level and the relationship between matter and energy by...**	
doing <u>both</u> of the following: • identifying chemical changes **AND** • identifying physical changes.	**doing <u>both</u> of the following:** • comparing the properties of original materials and their properties after undergoing chemical or physical change **AND** • observing and drawing conclusions about how the weight of an object compares to the sum of the weights of its parts.	**doing <u>both</u> of the following:** • explaining that all materials are made of small particles **AND** • identifying examples of chemical and physical changes.	**doing <u>both</u> of the following:** • explaining that adding heat causes the small particles in matter to move faster **AND** • demonstrating understanding that the properties of a material may change but the total amount of material remains the same.

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — D4
The Physical Setting – Force and Motion

Student understands the universal nature of matter, energy, force, and motion, and identifies how these relationships are exhibited in Earth Systems, in the solar system, and throughout the universe by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
summarizing how various forces affect the motion of objects by...			
doing the following: • identifying or demonstrating **one** way (e.g., forward, backward, straight, zigzag, up, down, fast, slow) an object can move.	**doing the following:** • identifying or demonstrating **two** ways (e.g., forward, backward, straight, zigzag, up, down, fast, slow) an object can move.	**doing <u>both</u> of the following:** • describing or demonstrating **three** ways (e.g., forward, backward, straight, zigzag, up, down, fast, slow) an object can move **AND** • identifying that the way an object moves can be changed by pushing or pulling it.	**doing the following:** • demonstrating understanding of how given objects move.

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
describing the force of gravity, the motion of objects, the properties of waves, and the wavelike property of energy in light waves by...		**See Extended Learning AGLEs**	**See Extended Learning AGLEs**
doing the following: • identifying or describing wave motions, earthquakes, vibrations, and/or water waves.	**doing <u>one</u> or more of the following:** • giving examples of how gravity pulls objects, • giving examples of how magnets pull and push objects, and/or • describing similarities in motion of sound vibration and earthquakes, and water waves.		

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — E1
The Living Environment — Biodiversity

Student understands that cells are the basic unit of life, that all life as we know it has evolved through genetic transfer and natural selection to create a great diversity of organisms, and that these organisms create interdependent webs through which matter and energy flow. Student understands the similarities and differences between humans and other organisms and the interconnections of these interdependent webs by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
comparing living things based on their behaviors, external features, and environmental needs by...			
doing the following: • identifying pictures or descriptions of given animals or plants.	**doing the following:** • identifying given organisms as plants or animals based on external features	**doing the following:** • identifying organisms that are similar and different based on external features, behaviors, and/or needs.	**doing <u>two</u> of the following:** • describing how plants and/or animals look, and/or • describing the things that plants and/or animals do, and/or • describing ways in which the needs of a plant and/or animal are met by its environment.

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
differentiating among organisms based on biological characteristics and identifying patterns of similarity by...		**describing and analyzing the evidence for relatedness among and within diverse populations of organisms and the importance of biodiversity by...**	
doing both of the following: • sorting living things based on external features or behaviors.	**doing <u>one</u> or more of the following:** • identifying how external (or internal) features can influence how an animal or plant gets food and/or • differentiating among living things that make their food, living things that eat their food, and those that do not clearly belong in one group or the other.	**doing <u>both</u> of the following:** • describing environments that have many different types of organisms and those that have fewer types of organisms, **AND** • identifying ways that organisms are related using physical evidence, such as presence or absence of a backbone.	**doing the following:** • predicting possible changes that could result if the numbers of different types of organisms were to be drastically reduced.

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — E2
The Living Environment — Ecosystems

Student understands that cells are the basic unit of life, that all life as we know it has evolved through genetic transfer and natural selection to create a great diversity of organisms, and that these organisms create interdependent webs through which matter and energy flow. Student understands the similarities and differences between humans and other organisms and the interconnections of these interdependent webs by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
describing ways organisms depend upon, interact within, and change the living and nonliving environment as well as ways the environment affects organisms by...			
doing the following: • identifying pictures or descriptions of given animals or plants.	**doing the following:** • identifying animals or plants that live in given environments	**doing the following:** • identifying plants, animals, and/or components of their environments in which given animals depend on for food and shelter.	**doing <u>one</u> of the following:** • comparing animals and plants that live in different environments to demonstrate understanding of how animals and plants depend on each other and the environments in which they live.

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
See Extended Learning AGLEs	**See Extended Learning AGLEs**	**See Extended Learning AGLEs**	**See Extended Learning AGLEs**

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — E3
The Living Environment — Cells

Student understands that cells are the basic unit of life, that all life as we know it has evolved through genetic transfer and natural selection to create a great diversity of organisms, and that these organisms create interdependent webs through which matter and energy flow. Student understands the similarities and differences between humans and other organisms and the interconnections of these interdependent webs by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
describing how living things are made up of one or more cells and the ways cells help organisms meet their basic needs by...			
doing the following: • identifying given parts of the human body.	**doing the following:** • matching animals and/or plants to their parts.	**doing the following:** • identifying parts that allow living things to meet basic needs.	**doing the following:** • identifying structures and/or processes that help given organisms stay alive.

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
describing the hierarchy of organization and function in organisms, and the similarities and differences in structure, function, and needs among and within organisms by...		**See Extended Learning AGLEs**	**See Extended Learning AGLEs**
doing <u>one</u> of the following: • identifying that some living things are made of one cell and some living things are made of many cells, and/or • identifying that all living things (single-celled and multi-celled) must have ways to get food and get rid of wastes.	**doing <u>both</u> of the following:** • identifying that some living things are made of one cell and some living things are made of many cells **AND** • identifying that all living things (single-celled and multi-celled) must have ways to get food and get rid of wastes.		

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — E4

The Living Environment — Hereditary and Reproduction

Student understands that cells are the basic unit of life, that all life as we know it has evolved through genetic transfer and natural selection to create a great diversity of organisms, and that these organisms create interdependent webs through which matter and energy flow. Student understands the similarities and differences between humans and other organisms and the interconnections of these interdependent webs by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
describing characteristics of organisms and the reason why organisms differ from or are similar to their parents by...			
doing the following: • identifying parents and their offspring by matching pictures of a baby organism to an adult of the same organism.	**doing the following:** • identifying things about offspring that are like and not like their parents.	**doing the following:** • demonstrating understanding of life cycles by explaining, drawing, or otherwise communicating knowledge of stages in given life cycles.	**doing <u>both</u> of the following:** • naming similarities between the adults and offspring of varied organisms **AND** • identifying and describing, drawing, or otherwise communicating knowledge of stages in a life cycle

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
describing the general characteristics and mechanisms of reproduction and heredity in organisms, including humans, and ways in which organisms are affected by their genetic traits by...		**See Extended Learning AGLEs**	**See Extended Learning AGLEs**
doing <u>one</u> of the following: • identifying the characteristics of offspring and parents based on similarities and differences.	**doing <u>both</u> of the following:** • identifying living things that reproduce by getting all their inherited information from one parent **AND** • identifying living things that reproduce by getting all their inherited information from two parents.		

Maine's Accountability Standards, Chapter 131 — Science AGLE/Indicator — E5
The Living Environment — Evolution

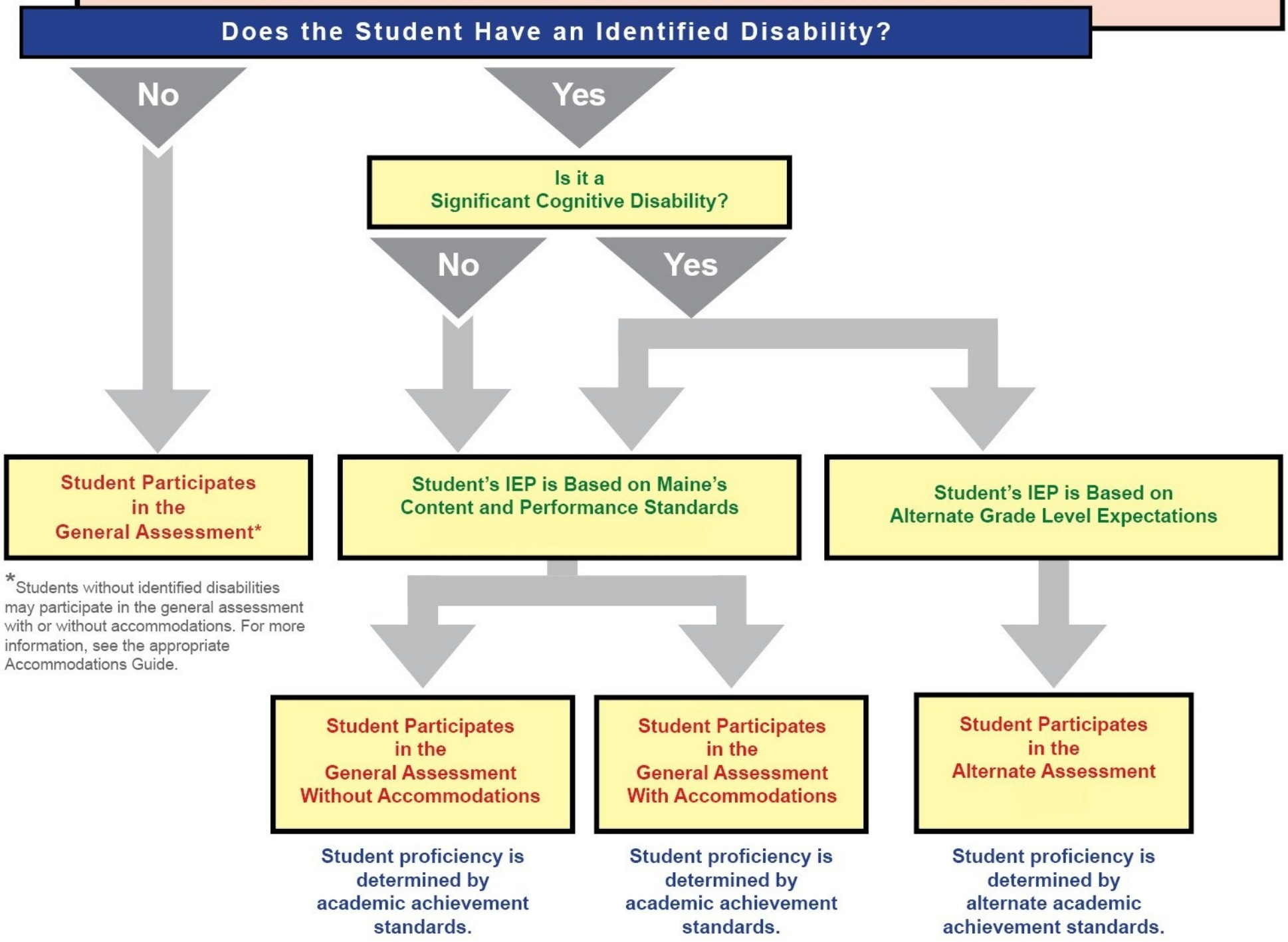
Student understands that cells are the basic unit of life, that all life as we know it has evolved through genetic transfer and natural selection to create a great diversity of organisms, and that these organisms create interdependent webs through which matter and energy flow. Student understands the similarities and differences between humans and other organisms and the interconnections of these interdependent webs by:

Level of Complexity 1	Level of Complexity 2	Level of Complexity 3	Level of Complexity 4
describing fossil evidence and present explanations that help us understand why there are differences among and between present and past organisms by...			
doing the following: • identifying organisms from the local environment.	**doing the following:** • matching pictures of organisms to the environment in which they live.	**doing <u>both</u> of the following:** • identifying organisms that no longer live today **AND** • describing features that organisms no longer living today share with organisms now alive and features that differ from those of organisms now alive.	**doing <u>both</u> of the following:** • describing features that allow or allowed present and past organisms to live in their environment **AND** • identifying organisms that once lived on Earth but no longer exist.

Level of Complexity 5	Level of Complexity 6	Level of Complexity 7	Level of Complexity 8
describing the evidence that evolution occurs over many generations, allowing species to acquire many of their unique characteristics or adaptations, by...		**describing the interactions between and among species, populations, and environments that lead to natural selection and evolution, by...**	
doing <u>both</u> of the following: • identifying examples of fossils **AND** • demonstrating understanding of how fossils are formed.	**doing the following:** • explaining how fossils are used to help us understand the past.	**doing the following:** • presenting explanations that help us understand similarities and differences among and between past and present organisms.	**doing <u>both</u> of the following:** • explaining why some organisms survive to the next generation **AND** • explaining why some organisms have traits that provide no apparent survival advantage.

APPENDIX B— PROCESS FOR DETERMINING THE APPROPRIATE AVENUE FOR PARTICIPATION

Flow Chart for Determining Appropriate Avenue of Assessment



APPENDIX C—2019 SCORING INSTRUCTIONS

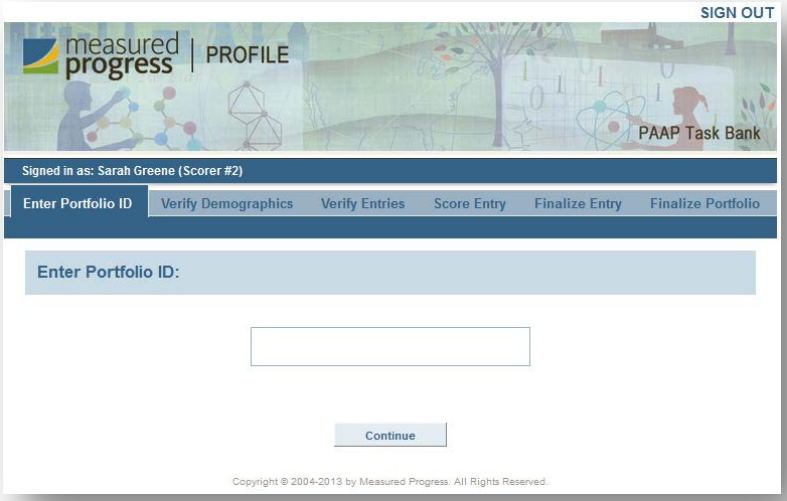
2019 PAAP

SCORING INSTRUCTIONS USING PROFILE

Step 1. Enter Portfolio ID

Step 1.a. Enter the 15-digit portfolio identification number (PID) found on the back of the portfolio envelope. This number begins with 753600000.

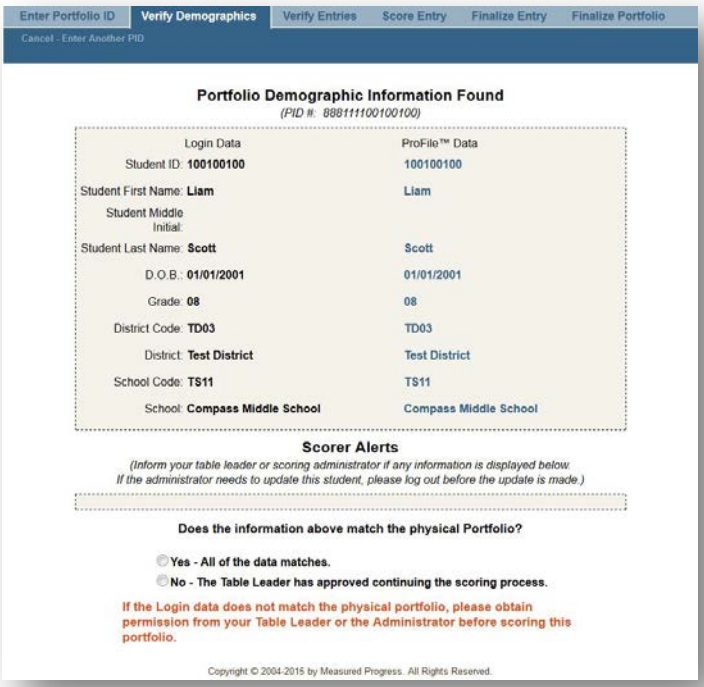
Step 1.b. Click **Continue**.



Step 2. Verify Demographics

Does the portfolio demographic information provided on the **Verify Demographics** screen match the login information on the portfolio? Compare the student ID number, name, grade, district name, and school name.

- If ***YES***, click **Yes** and then click on **Continue Scoring** in the dark blue banner.
- If ***NO***, click **No** notify your Table Leader. Once the Table Leader has approved, click on **Continue Scoring** in the dark blue banner.



NOTE: Navigate through ProFile by using the links within the application **ONLY**. Do **NOT** use the browsers back and forward buttons.

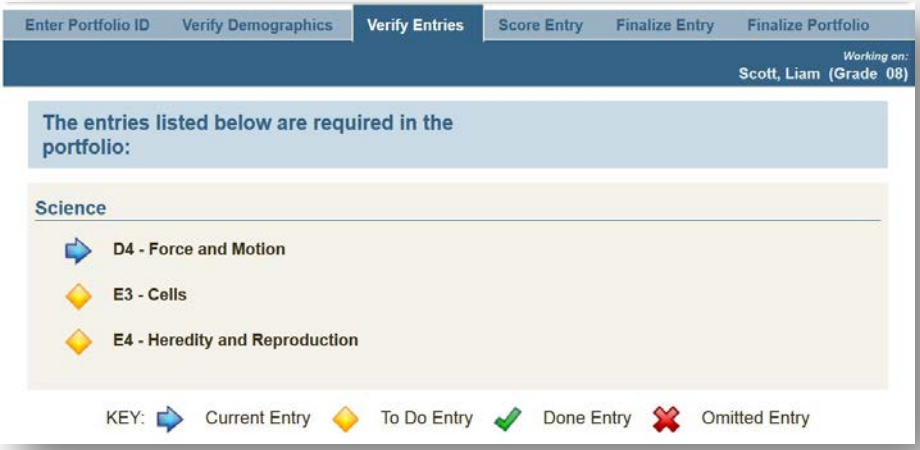
Step 3. Verify Entries

Step 3.a. Use the **Verify Entries** screen which lists the required entries and the Entry Slip to verify that the AGLE/Indicators on the screen match the circled AGLE/Indicator on the bottom of the Entry Slip.

In the example below, the AGLE/Indicators are D4, E3, & E4.

Science		
D.	Universe and Solar System	D1
	Earth	D2
	Matter and Energy	D3
	Force and Motion	D4

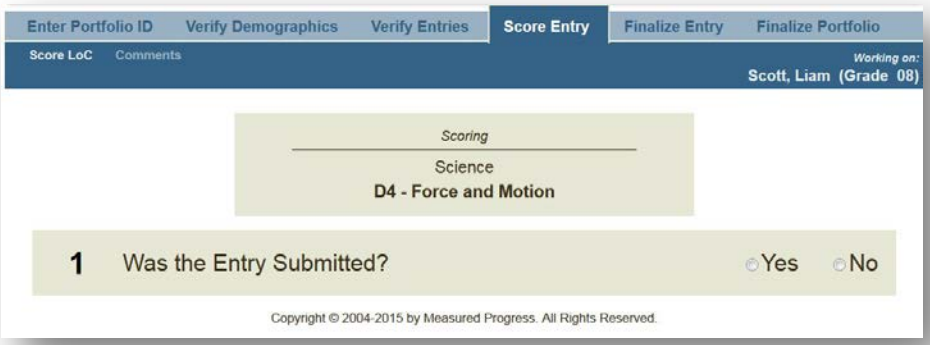
Step 3.b. Select the first Entry by clicking on the yellow diamond. Then a blue arrow will appear to indicate you have selected this Entry. Click the blue arrow to begin scoring the entry.



Step 4. Score Entry

Step 4.a. Does the content area submitted in the portfolio match what is shown on the screen? If you are not sure, check with your table leader before continuing.

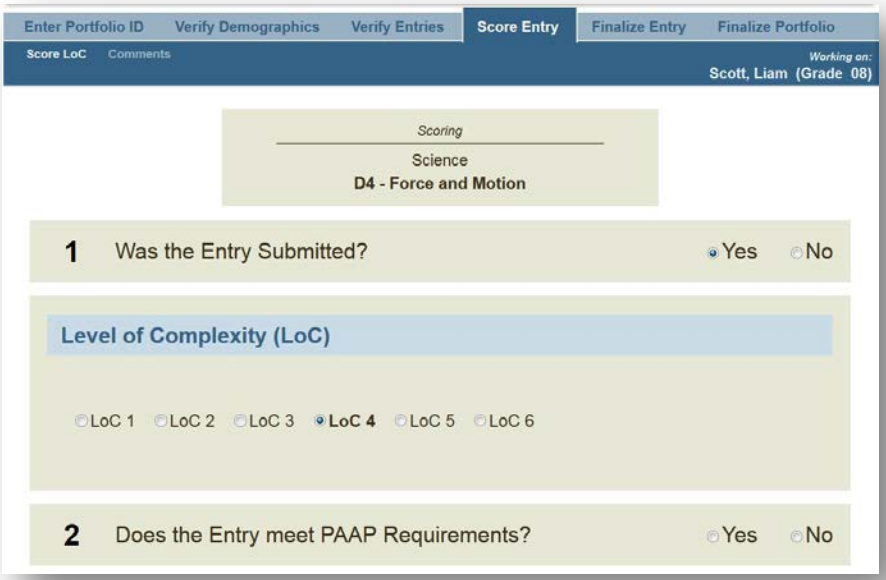
- If ***YES***, click **Yes** for "*Was the Entry Submitted?",* and continue scoring.
- If ***NO***, click **No** for "*Was the Entry Submitted?",* and click on **Comments** to assign comment code **2.e** and then **Finalize the Entry** (see page 8) before moving on to the next Entry.

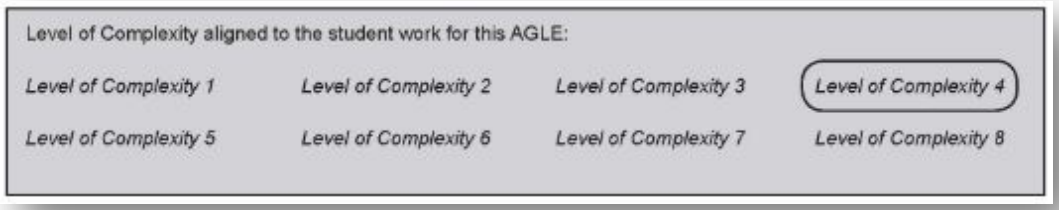


Step 4.b Level of Complexity (LoC)

The LoCs displayed on the screen are the only ones available for the student's grade level. Verify that the LoC circled in the middle of the Entry Slip page matches one of the LoCs on the screen. The LoC is also located on the bottom right corner of each page's footer within an Entry. In the examples below, the LoC is 4. Does the LoC circled on the Entry Slip match one of the LoCs on the screen?

- If ***YES***, select the LoC indicated on the Entry Slip or within the pages of the Entry.
- If ***NO***, then the Entry is unscorable and does not meet PAAP requirements. Do not enter anything for the LoC section and continue to **Step 4.c.**





Step 4.c. Does the Entry Meet PAAP Requirements? Verify that the LoC submitted is grade-appropriate. Only the grade-appropriate LoCs should appear on this screen. If you are not sure, check with your table leader before continuing.

- If ***YES***, click **Yes** for "*Does the Entry Meet PAAP Requirements?*"
- If ***NO***, click **No** for "*Does the Entry Meet PAAP Requirements?",* and click **Comments** in the dark blue banner. Assign comment code **4.a** or **4.b** and click **Finalize the Entry**.

Step 5. Score the Entry/Score Task X

Step 5.a. Is the Task Scorable?

Verify that the Task is scorable. **Both criteria below must be met.**

- There is evidence of student work on the work template for each task.
- The Level of Assistance was completed on the Task Summary page.
  - If "Other" was completed by the teacher, flag your table leader to verify that the Level of Assistance was selected accurately. Some of these issues are noted below.

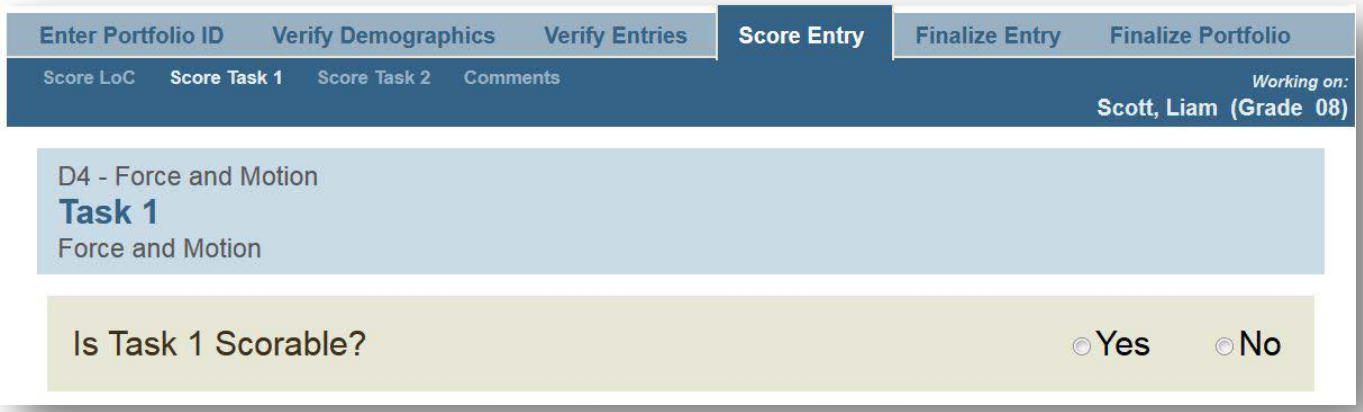
A Task is unscorable for any one of the following conditions:

- 3e - Student work was not completed on the work template.
- 5b - The Level of Assistance was not completed.
- 2c - Hand-over-hand was used.
- 2b - An item or items were altered.
- 3b - A Task or Task Summary page(s) is (are) missing.
- 4a - The LoC is above the student's grade level.

Special circumstances to consider:

- 4b - Two Entries for one AGLE/Indicator with different LoCs are submitted.
  - Score the Entry containing the latest date.
  - The other Entry is not scorable.

- If **YES**, click **Yes** for "*Is Task 1 Scorable?",* and continue scoring.
- If **No**, click **No** for "*Is Task 1 Scorable?",* and move to the next task by clicking on **Score Task X** in the dark blue banner.

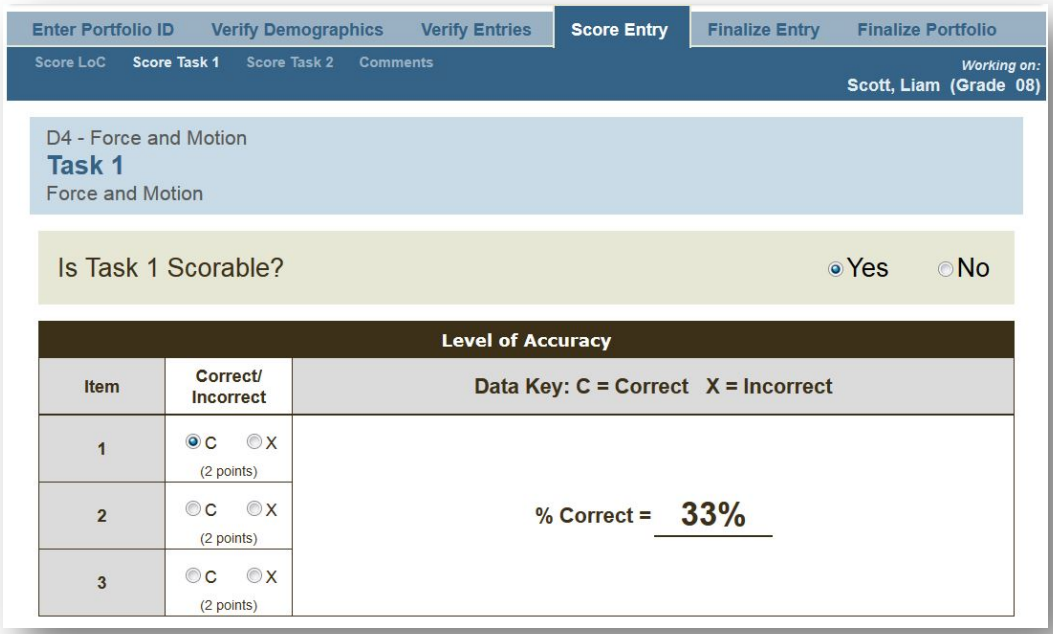


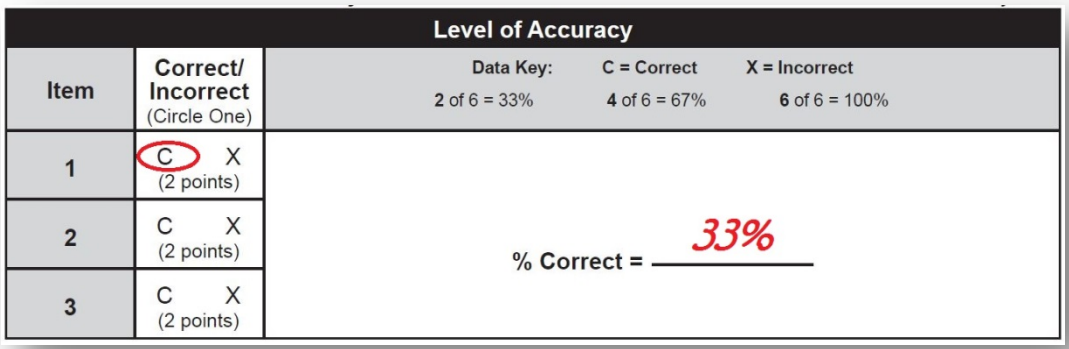
Step 5.b. Level of Accuracy

Use the *Level of Accuracy Grid* on the Task Summary page to identify the accuracy of the student work for each item. Refer to the "Responses Expected from Student" key on the Task Description page to score the Task.

- Click on **C** if the response is correct.
- Click on **X** if the response is incorrect.
- Verify that the "% Correct" increases when you click on **C**.

Note: If the percent correct reported by the teacher does not match your percent, do not change your score. Discrepancies will be handled by the third read process.

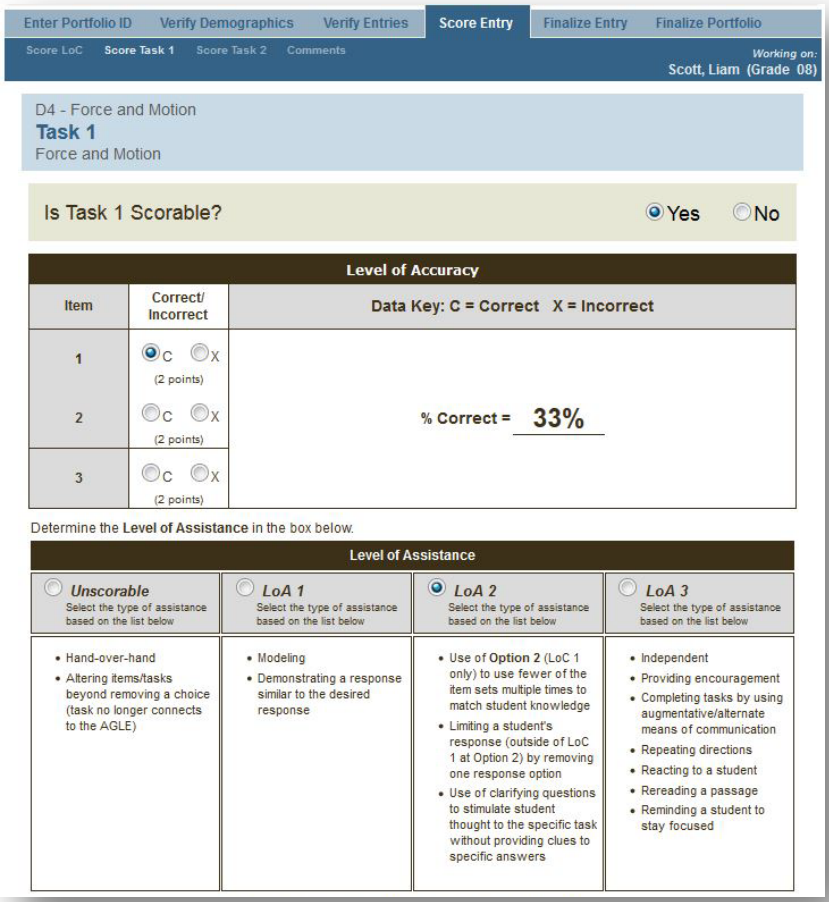


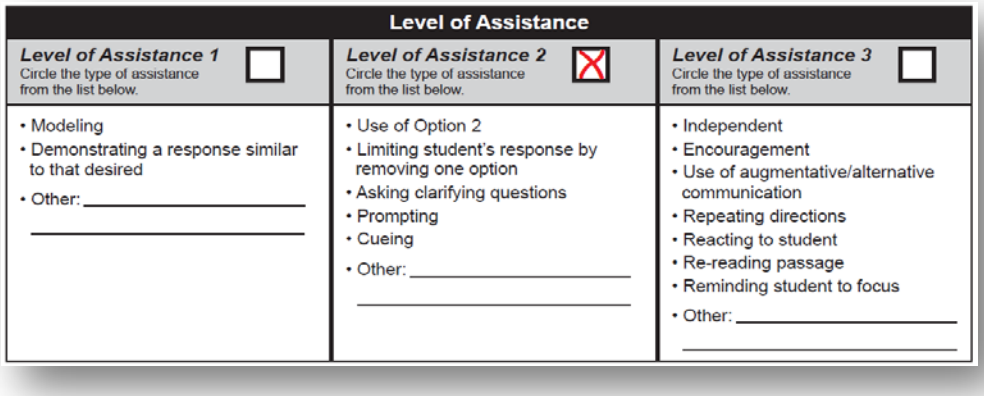


Step 5.c. Level of Assistance

Refer to the Task Summary page in the PAAP to complete the *Level of Assistance Grid*. If Applicable, compare the teacher's score and details provided under "Other" by the teacher to ensure there are no discrepancies. If you have a question about the "Other," check with your Table Leader before continuing. Is the Level of Assistance provided correct?

- If ***YES***, click the corresponding number 1, 2, or 3 in the Level of Assistance section of the screen.
- If ***NO***, flag your table leader to determine the correct Level of Assistance. After the revised Level of Assistance is determined, click this number in the Level of Assistance grid.
    - If the task is truly determined to be unscorable, return to the question at the top of the page "Is Task X Scorable?" and change your answer to "NO". Do not use the "Unscorable" option under Level of Assistance.



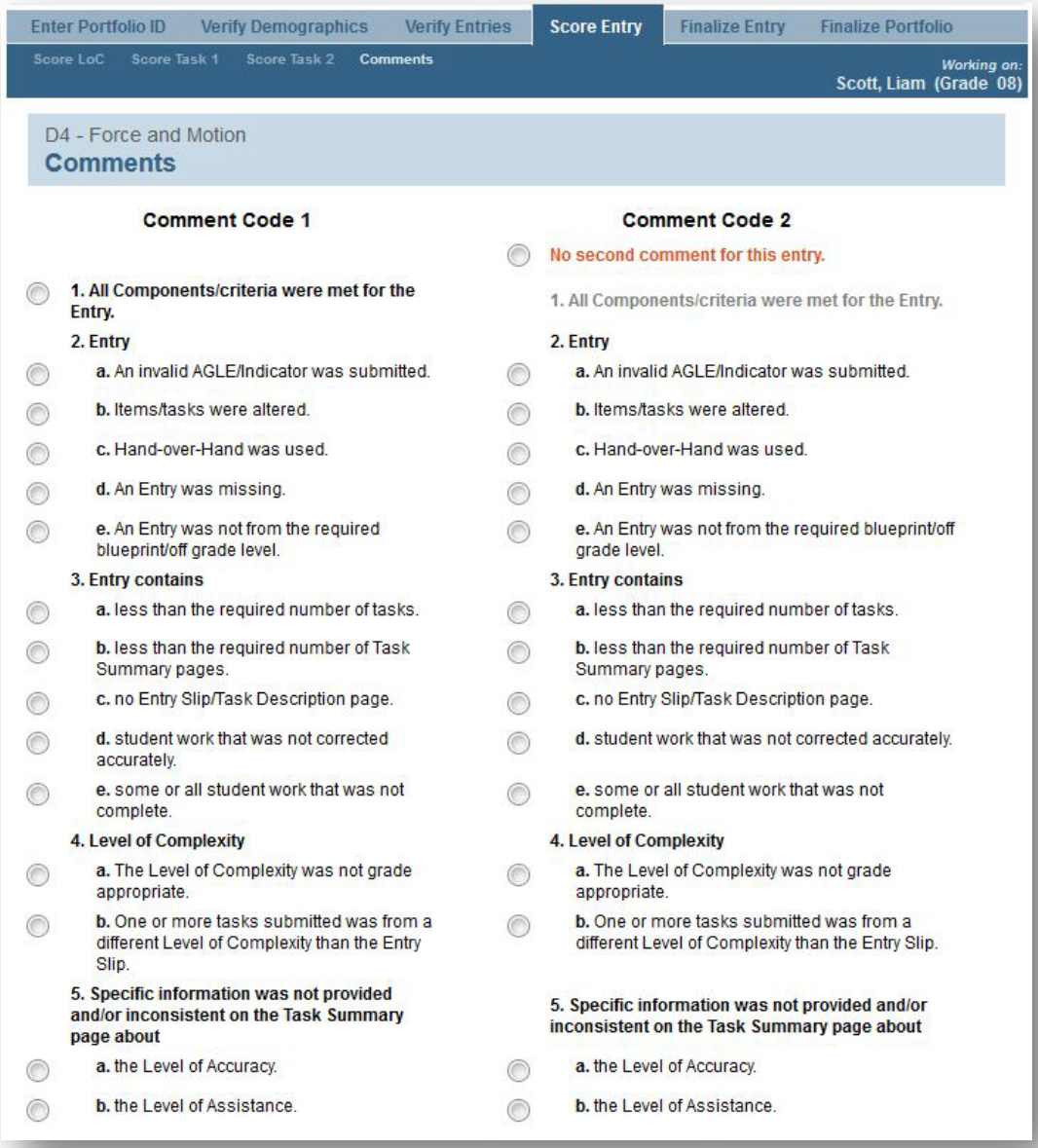


Step 5.d. Score the remaining Tasks for this Entry by clicking on **Score Task X** in the dark blue banner.

Step 5.e. Comment Codes

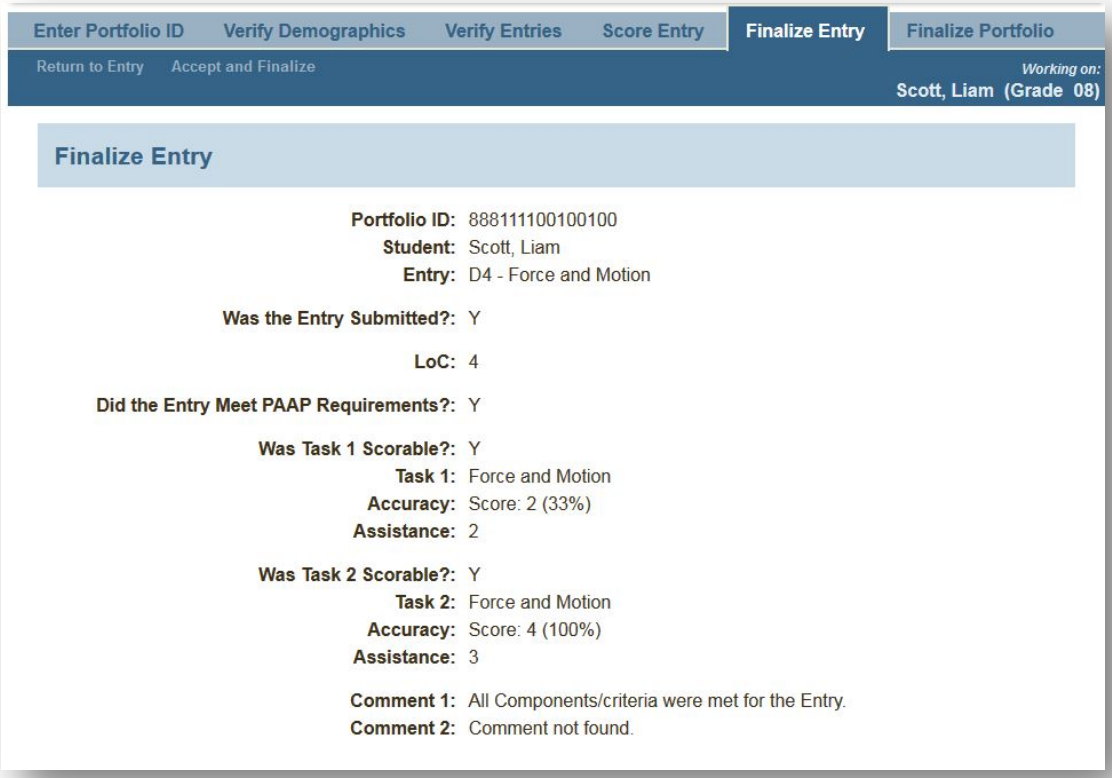
Comment codes are based on the totality of the Entry. They provide teachers valuable feedback on the Entry scores.

- Click on **Comments** in the dark blue banner.
- Select at least one comment code, but no more than two, as you score each Entry.



Step 5.f. When you have scored all the Tasks and selected appropriate comment codes, **Finalize Entry** will appear in the dark blue banner. If it does not appear, double check that all the Tasks were completed appropriately and comment codes were entered.

- Click on **Finalize Entry** and review the data that you entered for this Entry.
- If you notice an error, click on **Return to Entry** to verify that the data is correct for each Task. **Do <u>NOT</u> use the browsers back button.**
- Once it is determined that all data is accurate, click **Accept and Finalize**.



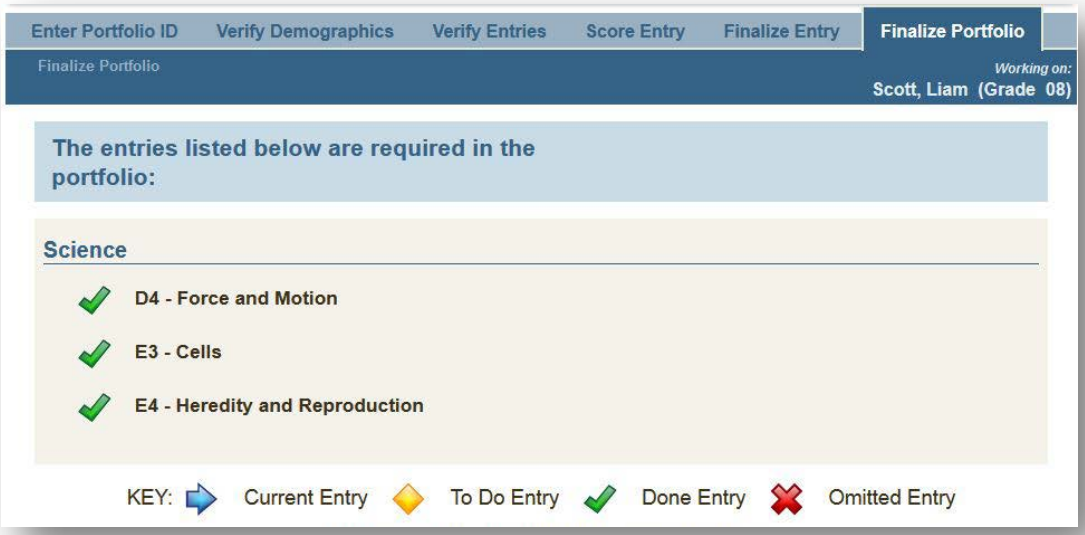
Note: Once you click **Accept and Finalize**, you

<u>CANNOT CHANGE OR REVIEW ANY DATA</u>.

Step 5.g. This Entry is now complete. Continue scoring the remaining Entries starting at **Step 3.a.** on page 2.

Step 6. Finalize the Portfolio

When you have scored all the Entries for the portfolio, ProFile will bring you to the **Finalize Portfolio** screen.



Verify that all Entries within the portfolio are completed in ProFile. If an Entry has not been completed, a yellow diamond will be displayed. Have all of the Entries been scored/reviewed?

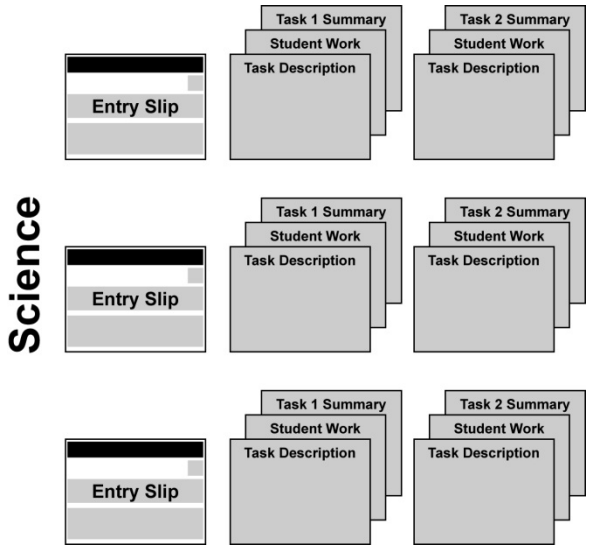
- If ***YES***, click **Finalize Portfolio** in the dark blue banner. You will then be prompted to enter the PID for the next portfolio to be scored.
- If ***NO***, flag your Table Leader.

Step 7. Flow of Materials

Once the scoring of the PAAP is complete:

- Place the PAAP back in the Tyvek envelope.
- Verify that you have indicated your scorer number in the proper place on the scoring label on the envelope.
- Return the PAAP to your table leader.

Visual Guide to the PAAP



ProFile Website

https://profile.measuredprogress.org/ProfileMEScoring/Login.aspx

Maine PAAP Blueprint

Grade	LoC	Science
5	1-4	D1, D2, E2
8	1-6	D4, E3, E4
3rd Yr. High School	1-8	D3, E1, E5

Comment Code	Scoring Issue/Irregularity	Scoring Rule
None	Entries (AGLE/Indicators) for a content area are not in alphabetical order (e.g., D4, E3, E4).	Flag your table leader who will reorganize the content area. Then you may continue scoring the PAAP.
1	**All components/criteria were met for the Entry.** Use only if **ALL** components/criteria were met for the Entry (all tasks).	Click on Comment Code 1 for the Entry.
2a	**An invalid AGLE/Indicator was submitted.** Repeated AGLE/Indicator across two Entries (e.g., D4, D4) **OR** AGLE/Indicator was not valid (e.g., D4 for a 3rd Yr HS student).	Flag your table leader. (Table leader should take these scenarios directly to the MDOE and Program Manager.) Click "No" for *Does the Entry meet PAAP Requirements*? for the entry that does not meet requirements. The MDOE and Program Manager will determine which Entry will not meet PAAP requirements on a case by case basis.
2b	**Tasks/items were altered.** Teacher read the passage/task to the student for an AGLE/Indicator/LoC requiring the student to read it (this is indicated in the Directions for Task Administration section of the Task Description page). **OR** Teacher changed the items/tasks in any way (e.g., the teacher replaced "experiences" with "objects" which no longer measures/assesses the same concept or skill.	For the task that was altered, click "No" for *Is Task X Scorable*?
2c	**Hand-over-Hand was used as a Level of Assistance for a task.**	For the task where Hand-over-Hand was used, click "No" for *Is Task X Scorable*?
2d	**An Entry is missing.** An Entry was not submitted (e.g., Reading requires two Entries and only one Entry was submitted).	Click "No" for *Was the Entry Submitted?*
3a	**An Entry contains fewer than the required number of tasks.** A task for an Entry is missing or incomplete.	If a task is missing from an Entry, click "No" for *Is Task X Scorable*? Score the other tasks submitted for that Entry if appropriate.
3b	**An Entry contains fewer than the required number of Task Summary pages.**	If a Task Summary page is missing, click "No" for *Is the Task Scorable?* Continue scoring the remaining tasks for that Entry.
3c	**An Entry is missing the Entry Slip and/or Task Description page(s).**	Continue scoring and provide this comment code (3c) for feedback to the teacher.
3d	**An Entry contains student work that was not corrected accurately by the teacher.**	Verify the expected student answers on the Task Description page. If the work was not correctly accurately, rescore the Level of Accuracy. Provide this comment code (3d) for feedback to the teacher.
3e	**An Entry contains incomplete student work.**	The completed tasks submitted may be scored if the Levels of Accuracy and Assistance were recorded on the Task Summary page. If all of the tasks contain incomplete work, click "No" for *Does the Entry meet PAAP Requirements*?
4a	**The LoC was not grade appropriate.**	If the grade assessed does not match the LoC for the Task submitted, verify the LoC with your table leader.
4b	**The LoCs for an Entry are not consistent.** One or more tasks submitted is/are from a different LoC than the Entry Slip.	If the LoC for one or more Tasks does not match the Entry Slip, flag your table leader. If all tasks are from the same LoC and only the Entry Slip is different, score the tasks from the same LoC that was selected and flag your table leader.
5a	**Specific information was not provided and/or inconsistent about the Level of Accuracy on the Task Summary page.**	If the student work is correct but the item(s) graded/corrected is not indicated on the Task Summary page, score the portfolio and flag your table leader. If the Level of Accuracy cannot be verified, click "No" for *Is the Task Scorable?*
5b	**Specific information was not provided and/or inconsistent on the Task Summary page about the Level of Assistance.**	If the Level of Assistance was selected but the specific assistance was not provided, flag your table leader. If the Level of Assistance was selected but it does not match the written description provided by the teacher, flag your table leader. If both the Level of Assistance and the description were not selected, click "No" for *Is the Task Scorable?*

APPENDIX D—SCORE OF RECORD

Maine Alt (PAAP) 1819 Score of Record

I. PAAP Portfolio-

All entries must be submitted with AGLE/Performance Indicators consistent with those listed in *2018-19_PAAP_Blueprint.pdf*. Each entry must have a unique AGLE/Performance Indictor.

1. Science: 3 entries submitted
   Grades 05, 08, 11 (3rd year HS)

All students must be reported in grade 05, 08, or 3rd year HS to be reported. Discrepancies will be resolved during data processing clean up.

II. Portfolio Data Points

Each portfolio will be scored at least twice. Some data points will require a third score. The scored data points are listed below.

1. PAAP Submitted: PAAP Submitted (Y, N)

2. Entry data points:
   a. Entry Submitted (Y, N, blank)
   b. AGLE (A, B, C, D, E, blank) (For non blank, see section I for valid values)
   c. Performance Indicators (1, 2, 3, 4, 5, blank) (For non blank, see section I for valid values)
   d. Level of Complexity (1-8, blank)
      i) Table of Valid Values

Grades	**Level of Complexity**
5	1, 2, 3, 4
8	1, 2, 3, 4, 5, 6
11	1, 2, 3, 4, 5, 6, 7, 8

   e. Entry Meets PAAP Requirements (Y, N, blank)

3. Task data points
   a. Scorable (Y, N, blank)

b. Level of Accuracy (1-4, blank)
c. Level of Assistance (1-3, blank)

4. Comment Codes (1, 2, 3, 4, 5, A, B, C, D, E, F)
   a. Valid value are 1, 2a, 2b, 2c, 2d,2e, 3a, 3b, 3c, 3d, 3e, 4a, 4b, 5a, 5b

III. Calculation of Final Score of Record for PAAP Submitted and Entry Data Points

1. Calculate Final PAAP Submitted
   a. If Scid_3 PAAP Submitted is not blank then Scid_3 PAAP Submitted is the Final PAAP Submitted. Else Scid_1 PAAP Submitted is the Final PAAP Submitted.
   b. If Final PAAP Submitted = 'N' then all entry data points are set to blank.
   c. If Final PAAP Submitted = 'Y' then calculate Final Entry data points and Comment Codes (as outlined below).

2. Calculate Final Entry Submitted, AGLE, Performance Indicator, Level of Complexity and Entry Meets PAAP Requirements
   a. If Scid_3 Entry Submitted is not blank then Scid_3 Entry Submitted is the Final Entry Submitted. Else Scid_1 Entry Submitted is the Final Entry Submitted.
   b. If Final Entry Submitted = 'N' then the AGLE, Performance Indicator, Level of Complexity, Entry Meets PAAP Requirements and all tasks data points are set to blank.
   c. If Final Entry Submitted = 'Y' then
      i. If Scid_3 AGLE is not blank then Scid_3 AGLE is the Final AGLE. Else Scid_1 AGLE is the Final AGLE.
      ii. If Scid_3 Performance Indicator is not blank then Scid_3 Performance Indicator is the Final Performance Indicator. Else Scid_1 Performance Indicator is the Final Performance Indicator.
      iii. If Scid_3 Level of Complexity is not blank then Scid_3 Level of Complexity is the Final Level of Complexity. Else Scid_1 Level of Complexity is the Final Level of Complexity.
      iv. If Scid_3 Entry Meets PAAP Requirements is not blank then Scid_3 Entry Meets PAAP Requirements is the Final Entry Meets PAAP Requirements. Else Scid_1 Entry Meets PAAP Requirements is the Final Entry Meets PAAP Requirements.
      v. If Final Entry Meets PAAP Requirements = 'N' then all task data points are set to blank.
      vi. If Final Entry Meets PAAP Requirements = 'Y' then for each task calculate Final Scorable.

3. Calculate Final Scorable, Level of Accuracy and Level of Assistance
   a. If Scid_3 Scorable is not blank then Scid_3 Scorable is the Final Scorable. Else Scid_1 Scorable is the Final Scorable.
   b. If Scorable = 'N' then the Level of Accuracy and Level of Assistance data points are set to 'U' (unscorable).
   c. If Scorable = 'Y' then

i. If Scid_3 Level of Accuracy is not blank then Scid_3 Level of Accuracy is the Final Level of Accuracy. Else Scid_1 Level of Accuracy is the Final Level of Accuracy.

d. If Scid_3 Level of Assistance is not blank then Scid_3 Level of Assistance is the final Level of Assistance. Else Scid_1 Level of Assistance is the Final Level of Assistance.

4. Calculate Final Comment Code(s)
   a. If Final PAAP Submitted = 'Y' then if Scid_3 Comment Code(s) is not blank then Scid_3 Comment Code(s) is/are the Final Comment code(s). Else Scid_1 Comment code(s) is/are the Final Comment Code.
   b. If Final PAAP Submitted = 'N' then set the Final Comment Codes to blank.

5. For entries with at least one scorable task, if the unique rule for AGLE/performance indicator described in section "I. PAAP Portfolio" is violated, then for the entry (entries) with the second (third) occurrence of the duplicate AGLE/performance indicator the final score of record must be Entry Submitted=Y, AGLE as calculated, Performance Indicator as calculated, Meets PAAP Requirements=N and all task data points must be blank.

IV. Calculation of Final Overall Achievement Scores based on Final Score of Record for PAAP Submitted and Entry Data Points

1. For when Final PAAP Submitted = Y a student will be assigned an Achievement Based Overall Score and Achievement Level.

2. For when Final PAAP Submitted = N, scores will be reported as No PAAP Submitted.

3. For each entry where Final Entry Submitted = N or Final Entry Meets Requirements = N or all tasks are unscorable then Final Entry Score = 0 and Final Entry Level of Accuracy (Assistance) = 'U'.

4. Final Entry Score = (5 * Final Level of Complexity) + Final Entry Level of Accuracy + Final Entry Level of Assistance – 4, where the following tables are used to calculate the Final Entry Scores for Level of Accuracy and Level of Assistance based on the number of tasks and total points across all tasks. For example, if an entry has 2 tasks (e.g. math) and the sum of the Level of Accuracy points across all tasks is 7 then the Final Entry Level of Accuracy score is 4. An unscorable task ('U') is assigned a score of 0 for calculation purposes.

Total Level of Accuracy Points

Number of Tasks	1	2	3	4	5	6	7	8	9	10	11	12
2	1	1	2	2	3	3	4	4				
3	1	1	1	1	2	2	2	3	3	3	4	4

Total Level of Assistance Points

Number of Tasks	1	2	3	4	5	6	7	8	9
2	1	1	2	2	3	3			
3	1	1	1	2	2	2	3	3	3

5. Overall Science Score = Sum of the Final Entry Scores

6. Overall Science Achievement Level will be determined based on the ranges of Overall Science Score. The ranges will be determined in standard setting and will be set by grade.

V. Valid Overall Science Score, Overall Science Achievement Level

1. For students identified as submitting a portfolio
    a. Overall Science Achievement Level: 1 (Substantially Below Proficient), 2 (Partially Proficient), 3 (Proficient), 4 (Proficient with distinction)
    b. Overall Science Score = 0 to max possible points which varies based on grade and subject

2. For students identified as not submitting a portfolio for Science
    a. No overall score. Student reported as No PAAP Submitted as detailed in decision rules.
3. Addenda 6/28: comment codes – use scorer 2, not scorer 1 comment codes

APPENDIX E—ITEM-LEVEL CLASSICAL STATISTICS

Table E-1. 2018–19 PAAP: Item-Level Classical Test Theory Statistics—Science Grade 5

Item Number	*Difficulty*	*Discrimination*
D111LAC	0.56	0.62
D111LAS	0.70	0.44
D112LAC	0.57	0.49
D112LAS	0.73	0.53
D121LAC	0.74	0.55
D121LAS	0.78	0.50
D122LAC	0.83	0.67
D122LAS	0.90	0.76
D131LAC	0.73	0.65
D131LAS	0.88	0.82
D132LAC	0.76	0.61
D132LAS	0.87	0.79
D141LAC	0.71	0.39
D141LAS	0.82	0.39
D142LAC	0.93	-0.13
D142LAS	0.82	0.37
D211LAC	0.47	0.61
D211LAS	0.68	0.78
D212LAC	0.43	0.61
D212LAS	0.62	0.79
D221LAC	0.77	0.66
D221LAS	0.83	0.74
D222LAC	0.78	0.72
D222LAS	0.87	0.80
D231LAC	0.78	0.46
D231LAS	0.91	0.38
D232LAC	0.80	0.43
D232LAS	0.91	0.32
D241LAC	0.85	0.28
D241LAS	0.80	0.34
D242LAC	0.80	0.43
D242LAS	0.78	0.49
E211LAC	0.71	0.65
E211LAS	0.76	0.61
E212LAC	0.64	0.74
E212LAS	0.75	0.61
E221LAC	0.77	0.75
E221LAS	0.84	0.58
E222LAC	0.68	0.57
E222LAS	0.84	0.59
E231LAC	0.71	0.37
E231LAS	0.90	0.62
E232LAC	0.77	0.41
E232LAS	0.93	0.66
E241LAC	0.84	0.29
E241LAS	0.79	0.28
E242LAC	0.84	0.42
E242LAS	0.82	0.37

Note: Statistics are presented only for items that were taken by 10 or more students.

Table E-2. 2018–19 PAAP: Item-Level Classical Test Theory Statistics—Science Grade 8

Item Number	*Difficulty*	*Discrimination*
D411LAC	0.45	0.93
D411LAS	0.53	0.87
D412LAC	0.38	0.91
D412LAS	0.53	0.87
D421LAC	0.71	0.43
D421LAS	0.58	0.55
D422LAC	0.60	0.49
D422LAS	0.58	0.66
D431LAC	0.61	0.68
D431LAS	0.73	0.72
D432LAC	0.75	0.92
D432LAS	0.73	0.76
D441LAC	0.76	0.70
D441LAS	0.74	0.58
D442LAC	0.76	0.70
D442LAS	0.77	0.61
D451LAC	0.85	0.55
D451LAS	0.87	0.17
D452LAC	0.86	0.11
D452LAS	0.85	0.33
D461LAC	0.83	0.53
D461LAS	0.90	0.07
D462LAC	0.77	0.79
D462LAS	0.88	0.19
E311LAC	0.29	0.48
E311LAS	0.42	0.55
E312LAC	0.33	0.41
E312LAS	0.50	0.27
E321LAC	0.78	0.84

continued

Item Number	Difficulty	Discrimination
E321LAS	0.79	0.79
E322LAC	0.79	0.73
E322LAS	0.81	0.73
E331LAC	0.77	0.82
E331LAS	0.82	0.87
E332LAC	0.79	0.90
E332LAS	0.75	0.82
E341LAC	0.74	0.51
E341LAS	0.85	0.51
E342LAC	0.77	0.76
E342LAS	0.90	0.59
E351LAC	0.83	0.30
E351LAS	0.91	0.27
E352LAC	0.85	0.18
E352LAS	0.93	0.22
E361LAC	0.73	0.38
E361LAS	0.96	-0.20
E362LAC	0.78	0.30
E362LAS	0.98	-0.05
E411LAC	0.44	0.85
E411LAS	0.50	0.86
E412LAC	0.45	0.71
E412LAS	0.48	0.72

Item Number	Difficulty	Discrimination
E421LAC	0.70	0.78
E421LAS	0.82	0.84
E422LAC	0.61	0.58
E422LAS	0.81	0.83
E431LAC	0.80	0.45
E431LAS	0.85	0.73
E432LAC	0.74	0.28
E432LAS	0.84	0.60
E441LAC	0.89	0.21
E441LAS	0.85	-0.08
E442LAC	0.82	-0.10
E442LAS	0.81	-0.14
E451LAC	0.92	-0.06
E451LAS	0.87	0.38
E452LAC	0.88	0.04
E452LAS	0.84	0.31
E461LAC	0.68	0.16
E461LAS	0.92	0.56
E462LAC	0.80	-0.16
E462LAS	0.97	0.32

Note: Statistics are presented only for items that were taken by 10 or more students.

Table E-3. 2018–19 PAAP: Item-Level Classical Test Theory Statistics—Science High School

Item Number	Difficulty	Discrimination
D311LAC	0.63	0.81
D311LAS	0.58	0.64
D312LAC	0.60	0.84
D312LAS	0.61	0.61
D321LAC	0.63	0.70
D321LAS	0.82	0.76
D322LAC	0.78	0.68
D322LAS	0.80	0.72
D331LAC	0.80	0.20
D331LAS	0.80	0.05
D332LAC	0.85	0.69
D332LAS	0.82	0.11
D341LAC	0.78	0.85
D341LAS	0.69	0.85
D342LAC	0.77	0.89
D342LAS	0.71	0.82
D351LAC	0.81	0.29
D351LAS	0.80	0.24
D352LAC	0.87	0.32
D352LAS	0.87	0.20
D361LAC		
D361LAS		

Item Number	Difficulty	Discrimination
D362LAC		
D362LAS		
D371LAC	0.77	-0.07
D371LAS	0.82	0.47
D372LAC	0.85	-0.01
D372LAS	0.77	0.02
D381LAC	0.90	0.78
D381LAS	0.97	-0.09
D382LAC	0.88	0.65
D382LAS	0.86	-0.21
E111LAC	0.68	0.93
E111LAS	0.76	0.89
E112LAC	0.70	0.80
E112LAS	0.76	0.89
E121LAC	0.75	0.90
E121LAS	0.63	0.83
E122LAC	0.65	0.63
E122LAS	0.57	0.63
E131LAC	0.89	0.63
E131LAS	0.85	0.52
E132LAC	0.78	0.42

continued

Item Number	Difficulty	Discrimination
E132LAS	0.78	0.54
E141LAC	0.60	0.76
E141LAS	0.62	0.92
E142LAC	0.73	0.85
E142LAS	0.64	0.92
E151LAC	0.82	-0.02
E151LAS	0.91	0.02
E152LAC	0.93	-0.18
E152LAS	0.93	0.04
E161LAC	0.80	0.08
E161LAS	0.88	-0.03
E162LAC	0.79	-0.04
E162LAS	0.88	0.03
E171LAC		
E171LAS		
E172LAC		
E172LAS		
E181LAC	0.85	0.90
E181LAS	0.87	0.79
E182LAC	0.82	0.89
E182LAS	0.87	0.78
E511LAC	0.73	0.78
E511LAS	0.81	0.81
E512LAC	0.75	0.76
E512LAS	0.79	0.74
E521LAC	0.81	0.63
E521LAS	0.69	0.52
E522LAC	0.82	0.77

Item Number	Difficulty	Discrimination
E522LAS	0.72	0.59
E531LAC	0.70	0.63
E531LAS	0.80	0.79
E532LAC	0.70	0.48
E532LAS	0.77	0.67
E541LAC	0.75	0.38
E541LAS	0.79	0.38
E542LAC	0.75	0.37
E542LAS	0.79	0.64
E551LAC	0.91	0.10
E551LAS	0.89	0.24
E552LAC	0.76	0.30
E552LAS	0.88	0.55
E561LAC		
E561LAS		
E562LAC		
E562LAS		
E571LAC	0.65	0.65
E571LAS	0.70	0.89
E572LAC	0.50	0.56
E572LAS	0.67	0.88
E581LAC	0.88	-0.11
E581LAS	0.77	0.37
E582LAC	0.88	0.42
E582LAS	0.73	0.48

Note: Statistics are presented only for items that were taken by 10 or more students.

APPENDIX F—ITEM-LEVEL SCORE DISTRIBUTION

Table F-1. 2018–19 PAAP: Item-Level Score Distributions for Constructed-Response Items—Science Grade 5

Item Number	*Total Possible Points*	*Percent of Students at Score Point*				
		0	*1*	*2*	*3*	*4*
D111LAC	4	7.41	11.11	44.44	22.22	14.81
D111LAS	3	7.41	18.52	29.63	44.44	
D112LAC	4	7.41	7.41	44.44	29.63	11.11
D112LAS	3	7.41	18.52	22.22	51.85	
D121LAC	4	16.95	1.69	8.47	15.25	57.63
D121LAS	3	16.95	5.08	5.08	72.88	
D122LAC	4	5.08	3.39	3.39	28.81	59.32
D122LAS	3	5.08	5.08	5.08	84.75	
D131LAC	4	3.33	3.33	30.00	23.33	40.00
D131LAS	3	3.33	3.33	20.00	73.33	
D132LAC	4	3.33	3.33	26.67	20.00	46.67
D132LAS	3	3.33	3.33	23.33	70.00	
D141LAC	4	0.00	10.53	36.84	10.53	42.11
D141LAS	3	0.00	15.79	21.05	63.16	
D142LAC	4	0.00	0.00	0.00	26.32	73.68
D142LAS	3	0.00	21.05	10.53	68.42	
D211LAC	4	8.70	34.78	26.09	21.74	8.70
D211LAS	3	8.70	17.39	34.78	39.13	
D212LAC	4	17.39	26.09	30.43	17.39	8.70
D212LAS	3	17.39	17.39	26.09	39.13	
D221LAC	4	6.45	6.45	3.23	41.94	41.94
D221LAS	3	6.45	6.45	19.35	67.74	
D222LAC	4	3.23	3.23	3.23	58.06	32.26
D222LAS	3	3.23	6.45	16.13	74.19	

Item Number	*Total Possible Points*	*Percent of Students at Score Point*				
		0	*1*	*2*	*3*	*4*
D231LAC	4	0.00	0.00	20.00	47.69	32.31
D231LAS	3	0.00	9.23	9.23	81.54	
D232LAC	4	0.00	1.54	13.85	47.69	36.92
D232LAS	3	0.00	7.69	10.77	81.54	
D241LAC	4	0.00	0.00	0.00	60.00	40.00
D241LAS	3	0.00	20.00	20.00	60.00	
D242LAC	4	0.00	6.67	13.33	33.33	46.67
D242LAS	3	0.00	20.00	26.67	53.33	
E211LAC	4	9.09	3.03	21.21	27.27	39.39
E211LAS	3	9.09	21.21	3.03	66.67	
E212LAC	4	9.09	6.06	27.27	36.36	21.21
E212LAS	3	9.09	21.21	6.06	63.64	
E221LAC	4	3.13	0.00	15.63	46.88	34.38
E221LAS	3	3.13	6.25	25.00	65.63	
E222LAC	4	3.13	3.13	37.50	31.25	25.00
E222LAS	3	3.13	6.25	25.00	65.63	
E231LAC	4	1.96	7.84	25.49	33.33	31.37
E231LAS	3	1.96	1.96	19.61	76.47	
E232LAC	4	1.96	1.96	31.37	15.69	49.02
E232LAS	3	1.96	1.96	11.76	84.31	
E241LAC	4	0.00	0.00	10.53	42.11	47.37
E241LAS	3	0.00	26.32	10.53	63.16	
E242LAC	4	0.00	0.00	15.79	31.58	52.63
E242LAS	3	0.00	15.79	21.05	63.16	

Table F-2. 2018–19 PAAP: Item-Level Score Distributions for Constructed-Response Items—Science Grade 8

Item Number	Total Possible Points	Percent of Students at Score Point				
		0	1	2	3	4
D411LAC	4	13.33	53.33	0.00	6.67	26.67
D411LAS	3	13.33	46.67	6.67	33.33	
D412LAC	4	13.33	60.00	6.67	0.00	20.00
D412LAS	3	13.33	46.67	6.67	33.33	
D421LAC	4	8.33	8.33	25.00	8.33	50.00
D421LAS	3	8.33	33.33	33.33	25.00	
D422LAC	4	8.33	33.33	8.33	8.33	41.67
D422LAS	3	8.33	41.67	16.67	33.33	
D431LAC	4	20.00	10.00	5.00	35.00	30.00
D431LAS	3	20.00	5.00	10.00	65.00	
D432LAC	4	15.00	0.00	5.00	30.00	50.00
D432LAS	3	15.00	10.00	15.00	60.00	
D441LAC	4	6.98	2.33	16.28	27.91	46.51
D441LAS	3	6.98	9.30	37.21	46.51	
D442LAC	4	6.98	4.65	11.63	32.56	44.19
D442LAS	3	6.98	6.98	34.88	51.16	
D451LAC	4	0.00	0.00	22.22	16.67	61.11
D451LAS	3	0.00	0.00	38.89	61.11	
D452LAC	4	0.00	5.56	5.56	27.78	61.11
D452LAS	3	0.00	5.56	33.33	61.11	
D461LAC	4	0.00	4.35	13.04	30.43	52.17
D461LAS	3	0.00	0.00	30.43	69.57	
D462LAC	4	4.35	4.35	21.74	17.39	52.17
D462LAS	3	4.35	0.00	21.74	73.91	
E311LAC	4	25.00	50.00	16.67	0.00	8.33
E311LAS	3	25.00	50.00	0.00	25.00	
E312LAC	4	16.67	58.33	8.33	8.33	8.33
E312LAS	3	16.67	50.00	0.00	33.33	
E321LAC	4	8.33	8.33	0.00	29.17	54.17
E321LAS	3	8.33	4.17	29.17	58.33	
E322LAC	4	8.33	8.33	0.00	25.00	58.33
E322LAS	3	8.33	0.00	33.33	58.33	
E331LAC	4	10.71	0.00	17.86	14.29	57.14
E331LAS	3	10.71	7.14	7.14	75.00	
E332LAC	4	10.71	0.00	14.29	14.29	60.71
E332LAS	3	10.71	10.71	21.43	57.14	
E341LAC	4	7.69	3.85	3.85	53.85	30.77
E341LAS	3	7.69	0.00	23.08	69.23	
E342LAC	4	3.85	0.00	11.54	53.85	30.77
E342LAS	3	3.85	0.00	19.23	76.92	
E351LAC	4	0.00	0.00	8.00	52.00	40.00
E351LAS	3	0.00	8.00	12.00	80.00	
E352LAC	4	0.00	0.00	16.00	28.00	56.00
E352LAS	3	0.00	8.00	4.00	88.00	
E361LAC	4	0.00	0.00	31.25	43.75	25.00
E361LAS	3	0.00	0.00	12.50	87.50	
E362LAC	4	0.00	0.00	12.50	62.50	25.00
E362LAS	3	0.00	0.00	6.25	93.75	
E411LAC	4	18.75	25.00	31.25	12.50	12.50
E411LAS	3	18.75	43.75	6.25	31.25	
E412LAC	4	25.00	18.75	18.75	25.00	12.50
E412LAS	3	25.00	31.25	18.75	25.00	
E421LAC	4	10.53	0.00	10.53	57.89	21.05
E421LAS	3	10.53	0.00	21.05	68.42	
E422LAC	4	10.53	5.26	31.58	36.84	15.79
E422LAS	3	10.53	0.00	26.32	63.16	
E431LAC	4	2.44	7.32	14.63	19.51	56.10
E431LAS	3	2.44	2.44	31.71	63.41	
E432LAC	4	4.88	9.76	17.07	21.95	46.34

continued

Item Number	Total Possible Points	Percent of Students at Score Point				
		0	1	2	3	4
E432LAS	3	4.88	4.88	24.39	65.85	
E441LAC	4	0.00	0.00	11.11	22.22	66.67
E441LAS	3	0.00	5.56	33.33	61.11	
E442LAC	4	5.56	0.00	11.11	27.78	55.56
E442LAS	3	5.56	5.56	27.78	61.11	
E451LAC	4	0.00	0.00	0.00	33.33	66.67
E451LAS	3	0.00	13.33	13.33	73.33	

Item Number	Total Possible Points	Percent of Students at Score Point				
		0	1	2	3	4
E452LAC	4	0.00	6.67	0.00	26.67	66.67
E452LAS	3	0.00	13.33	20.00	66.67	
E461LAC	4	4.55	4.55	13.64	68.18	9.09
E461LAS	3	4.55	0.00	9.09	86.36	
E462LAC	4	0.00	0.00	9.09	63.64	27.27
E462LAS	3	0.00	0.00	9.09	90.91	

Table F-3. 2018–19 PAAP: Item-Level Score Distributions for Constructed-Response Items—Science High School

Item Number	Total Possible Points	Percent of Students at Score Point				
		0	1	2	3	4
D311LAC	4	16.67	8.33	25.00	8.33	41.67
D311LAS	3	16.67	16.67	41.67	25.00	
D312LAC	4	16.67	8.33	25.00	16.67	33.33
D312LAS	3	16.67	16.67	33.33	33.33	
D321LAC	4	6.67	0.00	33.33	53.33	6.67
D321LAS	3	6.67	0.00	33.33	60.00	
D322LAC	4	6.67	0.00	13.33	33.33	46.67
D322LAS	3	6.67	6.67	26.67	60.00	
D331LAC	4	0.00	6.67	13.33	33.33	46.67
D331LAS	3	0.00	6.67	46.67	46.67	
D332LAC	4	0.00	6.67	13.33	13.33	66.67
D332LAS	3	0.00	6.67	40.00	53.33	
D341LAC	4	13.33	0.00	6.67	20.00	60.00
D341LAS	3	13.33	6.67	40.00	40.00	
D342LAC	4	13.33	0.00	6.67	26.67	53.33
D342LAS	3	13.33	6.67	33.33	46.67	
D351LAC	4	3.33	0.00	6.67	50.00	40.00
D351LAS	3	3.33	10.00	30.00	56.67	
D352LAC	4	0.00	3.33	6.67	30.00	60.00

Item Number	Total Possible Points	Percent of Students at Score Point				
		0	1	2	3	4
D352LAS	3	0.00	6.67	26.67	66.67	
D361LAC	4	12.50	0.00	12.50	37.50	37.50
D361LAS	3	12.50	0.00	37.50	50.00	
D362LAC	4	12.50	0.00	0.00	12.50	75.00
D362LAS	3	12.50	0.00	25.00	62.50	
D371LAC	4	0.00	7.69	7.69	53.85	30.77
D371LAS	3	0.00	0.00	53.85	46.15	
D372LAC	4	0.00	0.00	15.38	30.77	53.85
D372LAS	3	0.00	7.69	53.85	38.46	
D381LAC	4	0.00	0.00	8.33	25.00	66.67
D381LAS	3	0.00	0.00	8.33	91.67	
D382LAC	4	0.00	0.00	16.67	16.67	66.67
D382LAS	3	0.00	8.33	25.00	66.67	
E111LAC	4	9.09	9.09	9.09	45.45	27.27
E111LAS	3	9.09	9.09	27.27	54.55	
E112LAC	4	9.09	9.09	18.18	18.18	45.45
E112LAS	3	9.09	9.09	27.27	54.55	
E121LAC	4	10.00	0.00	20.00	20.00	50.00

continued

Item Number	Total Possible Points	Percent of Students at Score Point				
		0	1	2	3	4
E121LAS	3	10.00	20.00	40.00	30.00	
E122LAC	4	20.00	0.00	20.00	20.00	40.00
E122LAS	3	20.00	20.00	30.00	30.00	
E131LAC	4	5.00	0.00	5.00	15.00	75.00
E131LAS	3	5.00	0.00	30.00	65.00	
E132LAC	4	5.00	0.00	10.00	50.00	35.00
E132LAS	3	5.00	5.00	40.00	50.00	
E141LAC	4	15.38	7.69	30.77	15.38	30.77
E141LAS	3	15.38	0.00	69.23	15.38	
E142LAC	4	15.38	0.00	15.38	15.38	53.85
E142LAS	3	15.38	0.00	61.54	23.08	
E151LAC	4	0.00	0.00	5.56	61.11	33.33
E151LAS	3	0.00	5.56	16.67	77.78	
E152LAC	4	0.00	0.00	5.56	16.67	77.78
E152LAS	3	0.00	5.56	11.11	83.33	
E161LAC	4	0.00	0.00	11.54	57.69	30.77
E161LAS	3	0.00	3.85	26.92	69.23	
E162LAC	4	3.85	3.85	7.69	42.31	42.31
E162LAS	3	3.85	3.85	15.38	76.92	
E171LAC	4	14.29	0.00	0.00	71.43	14.29
E171LAS	3	14.29	0.00	71.43	14.29	
E172LAC	4	0.00	0.00	14.29	14.29	71.43
E172LAS	3	0.00	0.00	71.43	28.57	
E181LAC	4	6.67	0.00	6.67	20.00	66.67
E181LAS	3	6.67	0.00	20.00	73.33	
E182LAC	4	6.67	0.00	6.67	33.33	53.33
E182LAS	3	6.67	0.00	20.00	73.33	
E511LAC	4	7.14	7.14	14.29	28.57	42.86
E511LAS	3	7.14	7.14	21.43	64.29	
E512LAC	4	7.14	0.00	28.57	14.29	50.00

Item Number	Total Possible Points	Percent of Students at Score Point				
		0	1	2	3	4
E512LAS	3	7.14	7.14	28.57	57.14	
E521LAC	4	11.11	0.00	5.56	22.22	61.11
E521LAS	3	11.11	11.11	38.89	38.89	
E522LAC	4	5.56	0.00	11.11	27.78	55.56
E522LAS	3	5.56	11.11	44.44	38.89	
E531LAC	4	9.09	0.00	22.73	36.36	31.82
E531LAS	3	9.09	0.00	31.82	59.09	
E532LAC	4	13.64	4.55	4.55	40.91	36.36
E532LAS	3	13.64	0.00	27.27	59.09	
E541LAC	4	0.00	0.00	35.71	28.57	35.71
E541LAS	3	0.00	7.14	50.00	42.86	
E542LAC	4	0.00	0.00	28.57	42.86	28.57
E542LAS	3	0.00	21.43	21.43	57.14	
E551LAC	4	0.00	0.00	0.00	36.84	63.16
E551LAS	3	0.00	10.53	10.53	78.95	
E552LAC	4	0.00	0.00	36.84	21.05	42.11
E552LAS	3	0.00	5.26	26.32	68.42	
E561LAC	4	0.00	0.00	0.00	57.14	42.86
E561LAS	3	0.00	0.00	57.14	42.86	
E562LAC	4	14.29	0.00	42.86	28.57	14.29
E562LAS	3	14.29	14.29	28.57	42.86	
E571LAC	4	10.00	10.00	10.00	50.00	20.00
E571LAS	3	10.00	0.00	60.00	30.00	
E572LAC	4	10.00	30.00	10.00	50.00	0.00
E572LAS	3	10.00	0.00	70.00	20.00	
E581LAC	4	0.00	0.00	6.25	37.50	56.25
E581LAS	3	0.00	6.25	56.25	37.50	
E582LAC	4	0.00	0.00	6.25	37.50	56.25
E582LAS	3	0.00	25.00	31.25	43.75	

APPENDIX G—SUBGROUP RELIABILITY

Table G-1. 2018–19 PAAP: Subgroup Reliabilities—Science

Grade	*Group*	*Number of Students*	*Raw Score*			*Alpha*	*SEM*
			Maximum	*Mean*	*Standard Deviation*		
5	All Students	135	69	40.76	16.13	0.86	6.04
	Male	86	69	42.53	15.68	0.81	6.83
	Female	48	69	37.98	16.61	0.90	5.25
	Gender Not Reported	1	69				
	Hispanic or Latino	5	69				
	American Indian or Alaskan Native	6	69				
	Asian	0	69				
	Black or African American	5	69				
	Native Hawaiian or Pacific Islander	0	69				
	White (non-Hispanic)	111	69	40.77	16.81	0.88	5.82
	Two or more races	7	69				
	No Primary Race/Ethnicity Reported	1	69				
	Currently receiving LEP services	0	69				
	Former LEP student – monitoring year 1	0	69				
	Former LEP student – monitoring year 2	0	69				
	LEP: All Other Students	0	69				
	Students with an IEP	134	69	40.90	16.11	0.86	6.03
	IEP: All Other Students	1	69				
	Economically Disadvantaged Students	81	69	42.02	15.79	0.82	6.70
	SES: All Other Students	54	69	38.87	16.60	0.89	5.51
	Migrant Students	0	69				
	Migrant: All Other Students	135	69	40.76	16.13	0.86	6.04
	Students receiving Title 1 Services	0	69				
	Title 1: All Other Students	135	69	40.76	16.13	0.86	6.04
	Plan 504	0	69				
	Plan 504: All Other Students	135	69	40.76	16.13	0.86	6.04
8	All Students	131	99	57.60	26.88	0.89	8.92
	Male	90	99	56.58	27.02	0.90	8.54
	Female	41	99	59.85	26.76	0.89	8.88

continued

Grade	Group	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard Deviation		
	Gender Not Reported	0	99				
	Hispanic or Latino	9	99				
	American Indian or Alaskan Native	1	99				
	Asian	1	99				
	Black or African American	5	99				
	Native Hawaiian or Pacific Islander	1	99				
	White (non-Hispanic)	111	99	58.26	27.01	0.88	9.36
	Two or more races	3	99				
	No Primary Race/Ethnicity Reported	0	99				
	Currently receiving LEP services	0	99				
	Former LEP student – monitoring year 1	0	99				
8	Former LEP student – monitoring year 2	0	99				
	LEP: All Other Students	0	99				
	Students with an IEP	131	99	57.60	26.88	0.89	8.92
	IEP: All Other Students	0	99				
	Economically Disadvantaged Students	78	99	59.77	26.76	0.91	8.03
	SES: All Other Students	53	99	54.42	26.98	0.87	9.73
	Migrant Students	0	99				
	Migrant: All Other Students	131	99	57.60	26.88	0.89	8.92
	Students receiving Title 1 Services	0	99				
	Title 1: All Other Students	131	99	57.60	26.88	0.89	8.92
	Plan 504	0	99				
	Plan 504: All Other Students	131	99	57.60	26.88	0.89	8.92
	All Students	120	129	70.08	34.77	0.89	11.53
	Male	72	129	71.67	37.38	0.88	12.95
	Female	48	129	67.71	30.67	0.90	9.70
	Gender Not Reported	0	129				
HS	Hispanic or Latino	7	129				
	American Indian or Alaskan Native	3	129				
	Asian	1	129				
	Black or African American	1	129				
	Native Hawaiian or Pacific Islander	0	129				

continued

Grade	Group	Number of Students	Raw Score			Alpha	SEM
			Maximum	Mean	Standard Deviation		
HS	White (non-Hispanic)	106	129	69.71	34.64	0.88	12.00
	Two or more races	2	129				
	No Primary Race/Ethnicity Reported	0	129				
	Currently receiving LEP services	0	129				
	Former LEP student – monitoring year 1	0	129				
	Former LEP student – monitoring year 2	0	129				
	LEP: All Other Students	0	129				
	Students with an IEP	120	129	70.08	34.77	0.89	11.53
	IEP: All Other Students	0	129				
	Economically Disadvantaged Students	76	129	70.16	32.55	0.88	11.28
	SES: All Other Students	44	129	69.95	38.70	0.90	12.24
	Migrant Students	0	129				
	Migrant: All Other Students	120	129	70.08	34.77	0.89	11.53
	Students receiving Title 1 Services	0	129				
	Title 1: All Other Students	120	129	70.08	34.77	0.89	11.53
	Plan 504	3	129				
	Plan 504: All Other Students	117	129	69.65	34.79	0.89	11.54

APPENDIX H—DECISION ACCURACY AND CONSISTENCY RESULTS

Table H-1. 2018–19 PAAP: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—Overall and Conditional on Performance Level

Subject	*Grade*	*Overall*	*Kappa*	*Conditional on Level*			
				Level 1	*Level 2*	*Level 3*	*Level 4*
Science	5	0.77 (0.69)	0.54	0.85 (0.75)	0.69 (0.61)	0.81 (0.76)	0.76 (0.48)
	8	0.77 (0.68)	0.53	0.81 (0.70)	0.69 (0.58)	0.80 (0.75)	0.78 (0.58)
	HS	0.76 (0.68)	0.55	0.84 (0.76)	0.77 (0.69)	0.66 (0.61)	0.90 (0.68)

Table H-2. 2018–19 PAAP: Summary of Decision Accuracy (and Consistency) Results by Subject and Grade—Conditional on Cutpoint

Subject	*Grade*	*Not Proficient / Proficient*								
		Cut 1	*False*		*Cut 2*	*False*		*Cut 3*	*False*	
			Positive	*Negative*		*Positive*	*Negative*		*Positive*	*Negative*
Science	5	0.92 (0.90)	0.02	0.05	0.89 (0.84)	0.07	0.05	0.96 (0.95)	0.04	0.01
	8	0.93 (0.90)	0.03	0.03	0.89 (0.85)	0.05	0.05	0.94 (0.92)	0.04	0.02
	HS	0.91 (0.87)	0.05	0.05	0.91 (0.87)	0.04	0.05	0.94 (0.94)	0.06	0.00

APPENDIX I—INTERRATER CONSISTENCY

Table I-1. 2018–19 PAAP: Item-Level Interrater Consistency Statistics—Science Grade 5

Item	Number of		Percent		Correlation	Percent of Third Scores
	Score Categories	Responses Scored Twice	Exact	Adjacent		
D111LAC	4	25	84.00	12.00	0.85	16.00
D111LAS	3	25	92.00	4.00	0.85	8.00
D112LAC	4	25	84.00	8.00	0.71	16.00
D112LAS	3	25	92.00	4.00	0.86	8.00
D121LAC	4	49	100.00	0.00	1.00	2.04
D121LAS	3	49	97.96	0.00	0.87	4.08
D122LAC	4	56	89.29	8.93	0.82	14.29
D122LAS	3	56	98.21	0.00	0.88	3.57
D131LAC	4	28	100.00	0.00	1.00	0.00
D131LAS	3	28	96.43	0.00	0.80	3.57
D132LAC	4	28	100.00	0.00	1.00	0.00
D132LAS	3	28	96.43	0.00	0.65	3.57
D141LAC	4	19	57.89	26.32	0.65	42.11
D141LAS	3	19	100.00	0.00	1.00	0.00
D142LAC	4	19	89.47	10.53	0.84	10.53
D142LAS	3	19	100.00	0.00	1.00	0.00
D211LAC	4	19	89.47	10.53	0.96	10.53
D211LAS	3	19	94.74	5.26	0.96	5.26
D212LAC	4	17	88.24	11.76	0.95	11.76
D212LAS	3	17	100.00	0.00	1.00	0.00
D221LAC	4	28	96.43	3.57	0.98	3.57
D221LAS	3	28	96.43	0.00	0.84	3.57
D222LAC	4	29	100.00	0.00	1.00	0.00
D222LAS	3	29	96.55	0.00	0.83	3.45
D231LAC	4	64	93.75	4.69	0.89	6.25
D231LAS	3	64	98.44	1.56	0.98	1.56
D232LAC	4	64	93.75	4.69	0.91	6.25
D232LAS	3	64	98.44	1.56	0.98	1.56
D241LAC	4	15	86.67	13.33	0.82	13.33
D241LAS	3	15	100.00	0.00	1.00	0.00
D242LAC	4	15	80.00	20.00	0.90	20.00
D242LAS	3	15	100.00	0.00	1.00	0.00
E211LAC	4	29	100.00	0.00	1.00	0.00
E211LAS	3	29	100.00	0.00	1.00	0.00
E212LAC	4	29	100.00	0.00	1.00	0.00
E212LAS	3	29	100.00	0.00	1.00	0.00
E221LAC	4	31	96.77	3.23	0.97	6.45
E221LAS	3	31	93.55	0.00	0.63	6.45
E222LAC	4	31	100.00	0.00	1.00	0.00
E222LAS	3	31	96.77	0.00	0.85	3.23
E231LAC	4	49	63.27	10.20	0.37	36.73
E231LAS	3	49	97.96	2.04	0.95	2.04
E232LAC	4	48	70.83	6.25	0.49	29.17
E232LAS	3	48	95.83	2.08	0.64	4.17
E241LAC	4	18	66.67	22.22	0.33	33.33
E241LAS	3	18	100.00	0.00	1.00	0.00
E242LAC	4	18	50.00	50.00	0.10	50.00
E242LAS	3	18	94.44	5.56	0.95	5.56

Table I-2. 2018–19 PAAP: Item-Level Interrater Consistency Statistics—Science Grade 8

Item	Number of		Percent		Correlation	Percent of Third Scores
	Score Categories	Responses Scored Twice	Exact	Adjacent		
D411LAC	4	13	100.00	0.00	1.00	0.00
D411LAS	3	13	100.00	0.00	1.00	0.00
D412LAC	4	12	100.00	0.00	1.00	0.00
D412LAS	3	12	100.00	0.00	1.00	0.00
D421LAC	4	10	90.00	10.00	0.96	10.00
D421LAS	3	10	100.00	0.00	1.00	0.00
D422LAC	4	10	90.00	0.00	0.78	10.00
D422LAS	3	10	90.00	10.00	0.96	10.00
D431LAC	4	16	75.00	12.50	0.48	25.00
D431LAS	3	16	100.00	0.00	1.00	0.00
D432LAC	4	17	70.59	29.41	0.57	29.41
D432LAS	3	17	100.00	0.00	1.00	0.00
D441LAC	4	39	61.54	33.33	0.52	38.46
D441LAS	3	39	94.87	2.56	0.75	5.13
D442LAC	4	39	69.23	25.64	0.76	30.77
D442LAS	3	39	94.87	5.13	0.94	5.13
D451LAC	4	18	77.78	16.67	0.64	22.22
D451LAS	3	18	100.00	0.00	1.00	0.00
D452LAC	4	18	66.67	27.78	0.40	33.33
D452LAS	3	18	94.44	5.56	0.93	5.56
D461LAC	4	22	72.73	22.73	0.58	27.27
D461LAS	3	22	100.00	0.00	1.00	0.00
D462LAC	4	20	80.00	20.00	0.82	20.00
D462LAS	3	20	100.00	0.00	1.00	0.00
E311LAC	4	9				
E311LAS	3	9				
E312LAC	4	10	100.00	0.00	1.00	0.00
E312LAS	3	10	90.00	0.00	0.60	10.00
E321LAC	4	20	90.00	10.00	0.84	10.00
E321LAS	3	20	100.00	0.00	1.00	0.00
E322LAC	4	20	90.00	10.00	0.86	10.00
E322LAS	3	20	100.00	0.00	1.00	0.00
E331LAC	4	25	80.00	12.00	0.68	20.00
E331LAS	3	25	96.00	4.00	0.95	4.00
E332LAC	4	25	96.00	4.00	0.97	4.00
E332LAS	3	25	96.00	4.00	0.98	4.00
E341LAC	4	24	95.83	4.17	0.96	4.17
E341LAS	3	24	100.00	0.00	1.00	0.00
E342LAC	4	24	91.67	8.33	0.91	8.33
E342LAS	3	24	100.00	0.00	1.00	0.00
E351LAC	4	25	100.00	0.00	1.00	0.00

continued

Item	Number of		Percent		Correlation	Percent of Third Scores
	Score Categories	Responses Scored Twice	Exact	Adjacent		
E351LAS	3	25	100.00	0.00	1.00	0.00
E352LAC	4	25	96.00	4.00	0.97	4.00
E352LAS	3	25	100.00	0.00	1.00	0.00
E361LAC	4	15	80.00	20.00	0.85	20.00
E361LAS	3	15	100.00	0.00	1.00	0.00
E362LAC	4	15	66.67	33.33	0.63	33.33
E362LAS	3	15	100.00	0.00		0.00
E411LAC	4	13	84.62	15.38	0.93	15.38
E411LAS	3	13	84.62	7.69	0.81	15.38
E412LAC	4	12	91.67	8.33	0.96	8.33
E412LAS	3	12	100.00	0.00	1.00	0.00
E421LAC	4	17	94.12	5.88	0.92	5.88
E421LAS	3	17	100.00	0.00	1.00	0.00
E422LAC	4	17	94.12	5.88	0.96	5.88
E422LAS	3	17	100.00	0.00	1.00	0.00
E431LAC	4	40	90.00	5.00	0.78	10.00
E431LAS	3	40	97.50	2.50	0.96	2.50
E432LAC	4	39	97.44	2.56	0.99	2.56
E432LAS	3	39	97.44	0.00	0.75	2.56
E441LAC	4	18	88.89	11.11	0.88	11.11
E441LAS	3	18	100.00	0.00	1.00	0.00
E442LAC	4	17	64.71	29.41	0.14	35.29
E442LAS	3	17	100.00	0.00	1.00	0.00
E451LAC	4	14	64.29	21.43	0.27	35.71
E451LAS	3	14	100.00	0.00	1.00	0.00
E452LAC	4	14	78.57	0.00	0.20	21.43
E452LAS	3	14	100.00	0.00	1.00	0.00
E461LAC	4	20	95.00	5.00	0.95	5.00
E461LAS	3	20	100.00	0.00	1.00	0.00
E462LAC	4	21	95.24	4.76	0.94	4.76
E462LAS	3	21	100.00	0.00	1.00	0.00

Table I-3. 2018–19 PAAP: Item-Level Interrater Consistency Statistics—Science High School

Item	Number of		Percent		Correlation	Percent of Third Scores
	Score Categories	Responses Scored Twice	Exact	Adjacent		
D311LAC	4	10	100.00	0.00	1.00	0.00
D311LAS	3	10	90.00	10.00	0.92	10.00
D312LAC	4	10	100.00	0.00	1.00	0.00
D312LAS	3	10	100.00	0.00	1.00	0.00
D321LAC	4	14	100.00	0.00	1.00	0.00
D321LAS	3	14	100.00	0.00	1.00	0.00
D322LAC	4	14	92.86	0.00	0.76	7.14
D322LAS	3	14	100.00	0.00	1.00	0.00
D331LAC	4	14	85.71	14.29	0.87	21.43
D331LAS	3	14	100.00	0.00	1.00	7.14
D332LAC	4	13	92.31	0.00	0.70	15.38
D332LAS	3	13	100.00	0.00	1.00	7.69
D341LAC	4	13	76.92	15.38	0.28	23.08
D341LAS	3	13	100.00	0.00	1.00	0.00
D342LAC	4	13	76.92	15.38	0.54	23.08
D342LAS	3	13	100.00	0.00	1.00	0.00
D351LAC	4	29	89.66	10.34	0.85	10.34
D351LAS	3	29	100.00	0.00	1.00	0.00
D352LAC	4	30	76.67	20.00	0.73	23.33
D352LAS	3	30	100.00	0.00	1.00	0.00
D361LAC	4	7				
D361LAS	3	7				
D362LAC	4	7				
D362LAS	3	7				
D371LAC	4	13	100.00	0.00	1.00	0.00
D371LAS	3	13	100.00	0.00	1.00	0.00
D372LAC	4	13	92.31	7.69	0.93	7.69
D372LAS	3	13	100.00	0.00	1.00	0.00
D381LAC	4	12	100.00	0.00	1.00	0.00
D381LAS	3	12	100.00	0.00	1.00	0.00
D382LAC	4	12	91.67	8.33	0.93	8.33
D382LAS	3	12	100.00	0.00	1.00	0.00
E111LAC	4	10	90.00	10.00	0.95	10.00
E111LAS	3	10	100.00	0.00	1.00	0.00
E112LAC	4	10	100.00	0.00	1.00	0.00
E112LAS	3	10	100.00	0.00	1.00	0.00
E121LAC	4	9				
E121LAS	3	9				
E122LAC	4	8				
E122LAS	3	8				
E131LAC	4	19	100.00	0.00	1.00	0.00

continued

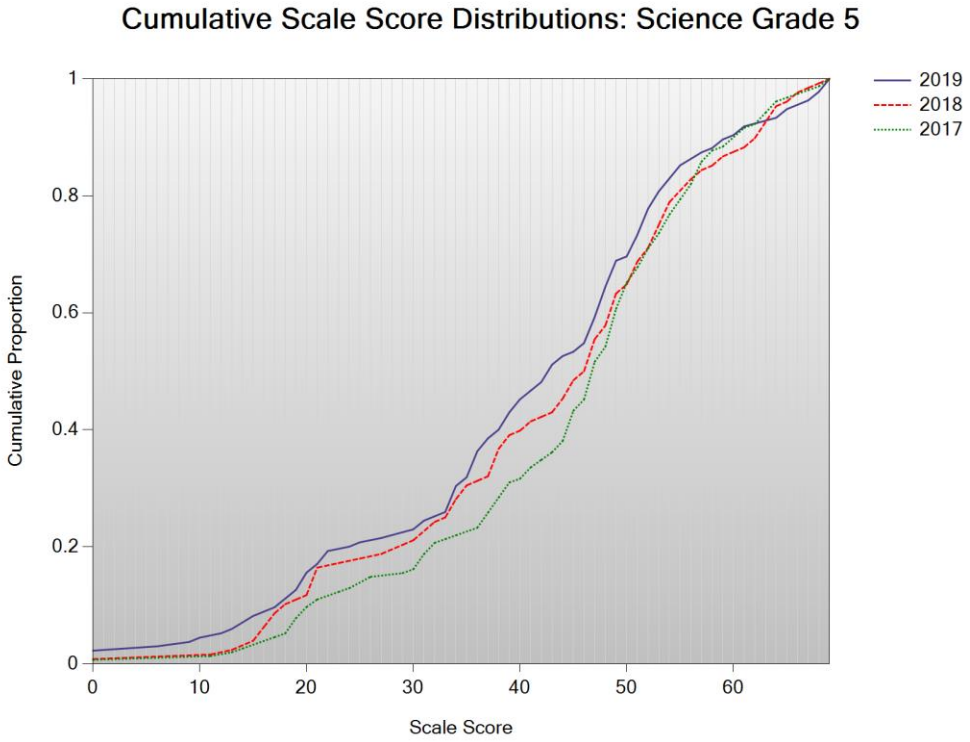
Item	Number of		Percent		Correlation	Percent of Third Scores
	Score Categories	Responses Scored Twice	Exact	Adjacent		
E131LAS	3	19	100.00	0.00	1.00	0.00
E132LAC	4	19	94.74	5.26	0.95	5.26
E132LAS	3	19	100.00	0.00	1.00	0.00
E141LAC	4	11	72.73	27.27	0.87	27.27
E141LAS	3	11	90.91	0.00	0.64	9.09
E142LAC	4	11	90.91	0.00	0.77	9.09
E142LAS	3	11	90.91	0.00	0.70	9.09
E151LAC	4	18	88.89	11.11	0.85	11.11
E151LAS	3	18	100.00	0.00	1.00	0.00
E152LAC	4	18	94.44	0.00	0.74	5.56
E152LAS	3	18	100.00	0.00	1.00	0.00
E161LAC	4	26	80.77	19.23	0.74	19.23
E161LAS	3	26	100.00	0.00	1.00	0.00
E162LAC	4	25	72.00	24.00	0.71	28.00
E162LAS	3	25	100.00	0.00	1.00	0.00
E171LAC	4	5				
E171LAS	3	5				
E172LAC	4	6				
E172LAS	3	6				
E181LAC	4	14	85.71	0.00	0.23	14.29
E181LAS	3	14	100.00	0.00	1.00	0.00
E182LAC	4	14	85.71	14.29	0.86	14.29
E182LAS	3	14	100.00	0.00	1.00	0.00
E511LAC	4	13	100.00	0.00	1.00	0.00
E511LAS	3	13	100.00	0.00	1.00	0.00
E512LAC	4	13	92.31	7.69	0.95	7.69
E512LAS	3	13	100.00	0.00	1.00	0.00
E521LAC	4	16	93.75	6.25	0.92	6.25
E521LAS	3	16	100.00	0.00	1.00	0.00
E522LAC	4	17	88.24	11.76	0.89	11.76
E522LAS	3	17	100.00	0.00	1.00	0.00
E531LAC	4	20	95.00	5.00	0.96	5.00
E531LAS	3	20	100.00	0.00	1.00	0.00
E532LAC	4	19	84.21	15.79	0.89	15.79
E532LAS	3	19	100.00	0.00	1.00	0.00
E541LAC	4	14	85.71	7.14	0.77	14.29
E541LAS	3	14	100.00	0.00	1.00	0.00
E542LAC	4	14	92.86	7.14	0.95	7.14
E542LAS	3	14	100.00	0.00	1.00	0.00
E551LAC	4	19	89.47	10.53	0.78	10.53
E551LAS	3	19	94.74	5.26	0.98	5.26
E552LAC	4	19	89.47	10.53	0.94	15.79
E552LAS	3	19	94.74	0.00	0.85	5.26

continued

Item	Number of		Percent		Correlation	Percent of Third Scores
	Score Categories	Responses Scored Twice	Exact	Adjacent		
E561LAC	4	7				
E561LAS	3	7				
E562LAC	4	6				
E562LAS	3	6				
E571LAC	4	7				
E571LAS	3	7				
E572LAC	4	7				
E572LAS	3	7				
E581LAC	4	16	81.25	12.50	0.48	18.75
E581LAS	3	16	100.00	0.00	1.00	0.00
E582LAC	4	16	87.50	6.25	0.73	12.50
E582LAS	3	16	93.75	6.25	0.95	6.25

APPENDIX J—CUMULATIVE SCORE DISTRIBUTIONS

Figure J-1. 2018–19 PAAP: Cumulative Distributions
Top: Science Grade 5 Bottom: Science Grade 8



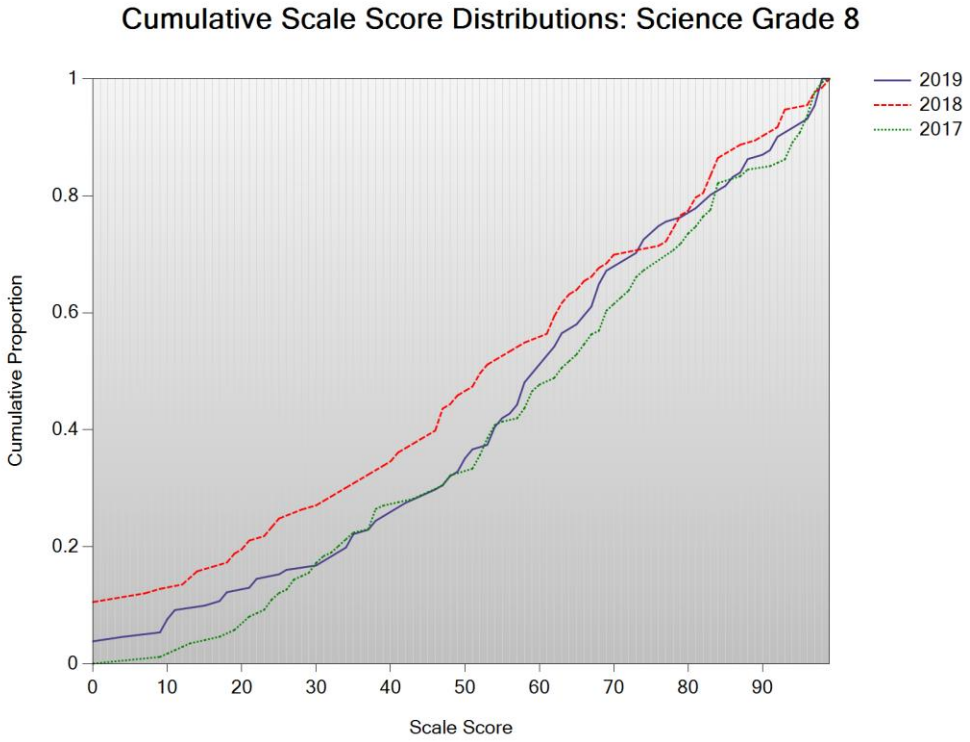
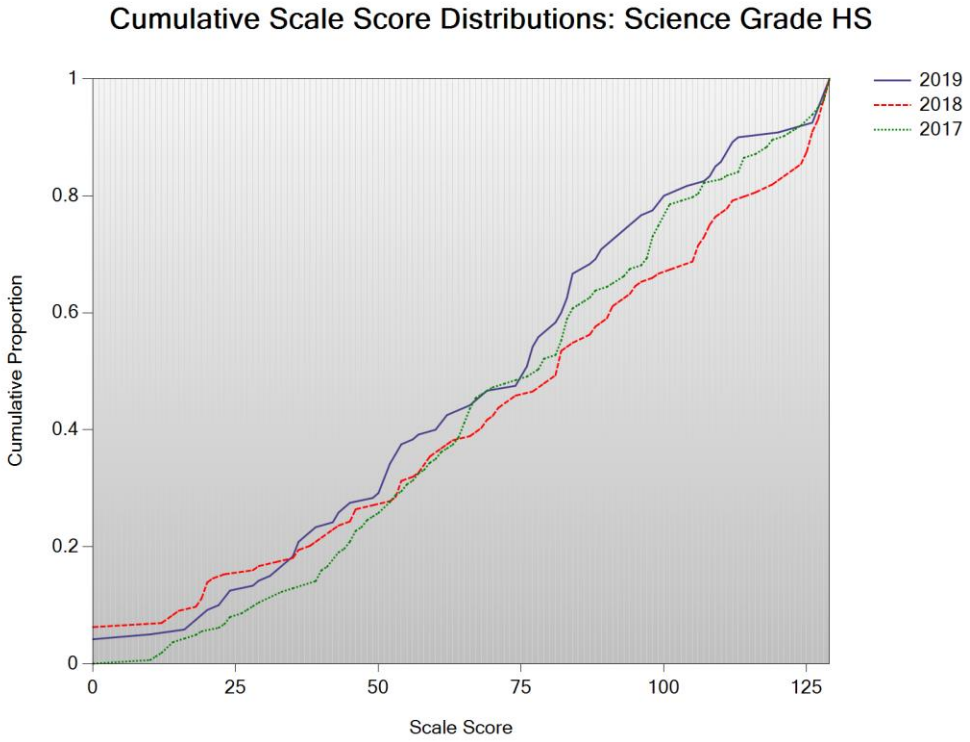


Figure J-2. 2018–19 PAAP: Cumulative Distributions
Top: Science High School



APPENDIX K—ACHIEVEMENT-LEVEL DISTRIBUTIONS

Table K-1. 2018–19 PAAP: Achievement-Level Distributions by Subject and Grade

Subject	*Grade*	*Achievement Level*	*Percent at Level*		
			2018–19	*2017–18*	*2016–17*
Science	5	1	19.26	17.19	12.26
		2	33.33	28.13	25.81
		3	42.22	50.78	58.71
		4	5.19	3.91	3.23
	8	1	16.79	28.57	18.97
		2	27.48	25.56	22.99
		3	45.80	37.59	43.68
		4	9.92	8.27	14.37
	HS	1	28.33	27.08	25.15
		2	38.33	27.78	36.81
		3	25.83	36.11	31.90
		4	7.50	9.03	6.13

APPENDIX L—PROCESSING AND REPORTING BUSINESS REQUIREMENTS

Processing and Reporting Business Requirements

Maine Education Assessment Personalized Alternate Assessment Portfolio			
753901: Spring 2019 Maine Alt (PAAP)			
Version Number	**Date**	**Updated Content Description**	**Updated By Name**
1.0	03/06/2019	Client Review Document	Keira Nevers
2.0	03/14/2019	Client Edits/Approval	Keira Nevers

Glossary

DOE	Department of Education
ELA	English Language Arts
PBT	Paper Based Test
CBT	Computer Based Test
LP	Large Print
TTS	Text to Speech
ASR	Assessed Student Review
PAAP	Personalized Alternate Assessment Portfolio
SRB	Student Scanned Booklet
SDF	Student Demographic File
SAU	School Administration Unit

Approval

I acknowledge that I have read this Decision Rules document and been informed of the contents of this document. By entering my name, title and date approved, I certify my approval. I have received a copy of this document for my records and understand that any further changes will require additional approvals as necessary.

Version	Printed Name	Title	Date Approved
2.0	Nancy Godfrey	Assessment Coordinator	03/25/2019

Table of Contents

I. Overview 3
A. Points of Contact 3
B. Document References 3
C. Assumptions 3
D. Risks 4
II. General Information 4
A. Assessments 4
B. Deliverables 4
C. Quality Assurance 4
III. IT Processing Pre-Test Assessment Administration 5
A. Student Roster and Test Data Preparation 5
B. School Type 5
C. Student Status 6
D. Other Information 6
IV. Post Test Assessment Administration 6
A. Assessed Student Review 7
B. Test Data 7
C. Scan Paper Delivery and Data Denotation 7
V. Participation and Exclusions 7
A. Test Attempt Rules 7
B. Student Participation Status 8
C. Student Participation Summary 8
VI. Calculations 8
A. Raw Scores 8
B. Scaling by Content Area 8
VII. Specific Datafile Rules 8
A. School Level Datafile (Summary) 8
B. State Student Overall Datafile 8
VIII. Datafile Table 9
IX. Non-Functional Requirements 9
A. Operational Requirements 9

I. Overview

This document will describe the Information Technology Processing and Reporting requirements for the 2018-2019 Maine PAAP testing window for the Maine Department of Education in support of providing assessment reporting for Maine PAAP final student test results.

A. Points of Contact

Title	Name	Contact Email
Client Services Program Manager	Jaimie Rissin	Rissin.Jaimie@measuredprogress.org
Project Management Functional Project Manager	Jimmy Hartman	Hartman.Jimmy@measuredprogress.org
Assessment Coordinator	Nancy Godfrey	Nancy.Godfrey@maine.gov
Assessment Data Analyst	Varun Motay	Varun.Motay@maine.gov
Project Manager Information Technology	Tony Scavuzzo	Scavuzzo.Tony@measuredprogress.org
Manager Information Technology Processing & Reporting	Sanjay Iyer	Iyer.Sanjay@measuredprogress.org
Manager Information Technology Software Quality Assurance	Scott Duquette	Duquette.Scott@measuredprogress.org
Senior Business Analyst Information Technology	Keira Nevers	Nevers.Keira@measuredprogress.org
Primary Processing Developer	Lezlie Turner	Turner.Lezlie@measuredprogress.org
Primary Data Analyst Engineer	Tyler Blouin	Blouin.Tyler@measuredprogress.org
Principal Data & Reporting Architect	Andrea Hebert	Hebert.Andrea@measuredprogress.org

B. Document References

Name the documents that were sources or results of this document that must be referenced to deliver the requirements as set forth in this document. Examples are File Layouts, and any referenced Technical Documents.

Pertains to the referenced documents and the location where the documents can be found. Each document shall be stored in the SharePoint site and have a Change Control audit log.

1) School Summary Layout
2) Student Results Layout
3) Level of Complexity Layout
4) State Student Scored Data Layout

C. Assumptions

In order to commit to delivering Data and Printed reports to the client, the following assumptions must be assumed.

D. Risks

Any risks shall be identified and recorded in their respective repositories. All stakeholders shall be notified of any risks associated to their responsible area's and be engaged as necessary.

1) Reporting issues from last Administration shall be mitigated
   a) The Student Demographic File was transmitted from DOE to MP IT Processing with incorrect values for Active Status field
      i) The DOE expected Labels for students, yet were not indicated as Active value '2'
      ii) Grade 11 students are considered Third Year High School and shall be considered for producing Labels, yet shall not be considered in Aggregate Reporting

Maine Student Demographic Layout				
20	Active	1	Active Status	1=Yes, 0=No, 2 = 3rd Yr HS

II. General Information

A. Assessments

1) Portfolio Assessment

Contract	Test Administration Open	Test Administration End
Science Grades 5, 8 & 3rd Year HS	March 1, 2019	March 30, 2019

B. Deliverables

Data Files delivered shall pass all validation rules and formats based on the layout and specification document.

1) Maine Alt Science
   a) School Level Data (Summary)
   b) State Student Overall Data
   c) State Student Entry Scores
   d) State Level of Complexity Data

C. Quality Assurance

All datafiles and reports identified as a deliverable to the Client shall pass internal Quality Assurance measures. The Software Quality Assurance (SQA) team works together with the data processing and data analysis teams to ensure quality data is captured and delivered accurately. Quality control checks are being performed by the data processors and data analysts as the data is handed off via multiple internal software tools. Included in the final execution, the Software Quality Assurance team executes test cases validating student printed reports and student labels for accuracy in comparison to the previously agreed upon report design specifications.

III. IT Processing Pre-Test Assessment Administration

Pre-Test Assessment Administration activities shall be completed prior to the Test Assessment Administration window. The Pre-Administration window shall allow for the client to gather the student and testing subject data to provide Cognia and all other vendors the information to administer the Test Assessments.

A. Student Roster and Test Data Preparation

1) MP shall receive the State Student Demographic file with all students from the DOE in the specified file format and layout (Maine Student Demographic Layout.xlsx)
   a) The incoming Student Demographic File shall be comma separated value file format (csv)
   b) The incoming Student Demographic File shall contain a Header row
   c) The incoming Student Demographic File shall use a comma to delimit data fields, and shall not contain any other delimiter values
   d) The Pre-ID file shall indicate any paper booklets and any accommodated booklets using the TestType field
   e) The Pre-ID file shall be validated upon receipt
      i) Validate all school and district codes exist in MP's iCore Database
      ii) Date of Birth values must be numeric, else MP will set to blank
      iii) First Name or Last Name fields shall contain no special characters, MP will set special characters to blank
      iv) Student ID's must be unique across all Schools and Districts
      v) MP shall compare Student ID's with prior year and report all name differences
      vi) MP shall compare Student ID's with prior year grade and report all students that remained the same grade
      vii) MP shall report the count of students by grade and test type

B. School Type

SchType	**Source:** **ICORE SubTypeID**	**Description**
'PUB'	1	Public
'PSP'	19	Public Special Purpose
'PSE'	15	Public Special Ed
'BIG'	6	Private with 60% or more Publicly Funded (Big 11)
'PSN'	23	Special Purpose Private
'CHA'	11	Public Charter

C. Student Status

Level	Impact on Analysis	Impact on Reporting
Student	n/a	School and SAU data will be blank for students with a StuStatus value of 1. Always print tested year state data. For StuStatus values of 1 School name is 'Home Schooled' and SAU name is the name of the student's reported SAU.
School	Exclude all students with a StuStatus value of 1 or 2.	n/a
SAU	Exclude all students with a StuStatus value of 1 or 2.	n/a
State	Exclude all students with a StuStatus value of 1, 2,, 4.	n/a.

D. Other Information

1) Public School districts are districts containing at least one school with a school sub-type-id of 1, 11, 15, or 19.
2) Student Demographic File Linking
   a. All alternately assessed students link to the Student Demographic File.
   b. All demographic data of record are pulled from the Student Demographic File for alternately assessed students
   c. Students are included in school and SAU aggregations, but not state aggregations
   d. Students will receive an ISR and will be listed on the school analysis report.
3) Third Year HS
   a. The Student Demographic File Grade is the student's grade used in reporting. Students identified as Third Year HS (Active = '2') will be treated as Third Year HS regardless of grade
   b. Third Year HS students are stored internally as Grade 11.
4) Only students in 5, 8, and Third Year HS are expected to test Science in 17-18. Data processing will provide discrepancy reports for all students submitting a Science PAAP that are not identified as grade 5, 8, or Third Year HS for resolution
5) Only students in 5, 8, and Third Year HS after resolution/clean up with a submitted PAAP are included in the Maine Alt reporting. Students that do not submit a PAAP are reported through Maine Science

IV. Post Test Assessment Administration

The Test Assessment Administration window shall be defined and closed prior to processing and reporting for Reporting. The commencement of the testing window shall initiate activity to complete all Results and Reporting to the Client.

A. Assessed Student Review

The client shall be provided a Post Test Clean Up five-day time frame which shall be used for additional post-test clean-up of Student Participation and Demographic record modifications

2) A specified window shall be determined where the client shall have the opportunity to update all allowable edits necessary to the Post Test records
3) DOE shall use the ASR file layout and specifications during the clean-up window
   a) The ASR File Specification has specific fields identified that may be edited as well as not
   b) Any field that does not allow edits will be ignored during processing

B. Test Data

1) Every imported Test record shall have an associated Student record
2) Test Mode shall be captured in all test records as 2: Paper
3) Paper Test Items shall have an associated BarcodeID assigned with the initial Student Demographic file which Cognia will use to link the Test metadata including, but not limited to Test Item, Grade, Subject and Session
4) Duplicate test records shall be merged/resolved prior to processing
   a) All duplicate records shall be reviewed and updated accordingly based on SDE feedback as necessary
   b) In the event the Student has a Test record with no items attempted, Cognia shall not suppress any records unless specifically directed to do so as part of the duplicate resolution

C. Scan Paper Delivery and Data Denotation

Each Paper Booklet test shall be scanned and delivered immediately to IT Processing team. At the time of receipt, IT Processing shall perform procedures to accurately identify inaccuracies in the data. IT Processing shall receive final raw scored data at time of delivery.

1) The data file shall be formatted as specified in the Scan Delivery Layout Format
2) Any and all discrepancies with the Scan File shall be resolved accordingly
3) IT Processing shall validate each field for each Student, Test and Item metadata to ensure the data is acceptable
4) IT Processing shall provide a report of any and all discrepancies back to the Scanning department for research and/or re-scanning
5) IT Processing shall validate each Label Barcode and update as necessary

V. Participation and Exclusions

A. Test Attempt Rules

1) Attempt PAAP: Participated in PAAP
   a) All students included in the Data Processing views for PAAP reporting have met the requirements for participation in PAAP Science for 17-18. See the Maine Alt Data Processing Specifications and the MaineAlt1718ScoreofRecord.pdf for details.

B. Student Participation Status

1) Tested
   a) Incomplete Portfolio: a required entry was submitted, but at least one required entry was not submitted
   b) Complete Portfolio: all required entries were submitted

C. Student Participation Summary

Participation Status	Part. Flag	Raw Score	Ach. Level
Tested: Alternate Assessment	C	✓	✓

VI. Calculations

Calculations shall be defined for reporting purposes.

A. Raw Scores

1) Refer to MaineAlt1819ScoreOfRecord.pdf

B. Scaling by Content Area

1) Achievement levels are assigned using a look-up table based on the student's raw score and grade

VII. Specific Datafile Rules

A. School Level Datafile (Summary)

2) State level csv file containing the number of students tested, and the number and percent of tested students at each achievement level; aggregated to the school level

B. State Student Overall Datafile

1) State level csv file containing student demographic data and performance information
2) Only students from 'PUB', 'CHA', 'PSP', and 'PSE' schools are included, or if they have a sending SAU

VIII. Datafile Table

(YYYY indicates School Year)

File	Delivery	Layout	Naming Convention
School Level Data File (Summary)	State	MaineAlt*YYYY*SchoolSummaryLayout.xls	MaineAlt*YYYY*SchoolSummaryData.csv
State Student Overall Data	State	MaineAlt*YYYY*StateStudentScoredDataLayout.xls (Worksheet: "Overall")	MaineAlt*YYYY*StateStudentScoredData.csv
State Student Entry Scores	State	MaineAlt*YYYY*StateStudentScoredDataLayout.xls (Worksheet: "EntryScores")	MaineAlt*YYYY*StateStudentEntryScoresData.csv
State Level of Complexity Data	State	MaineAlt*YYYY*LOCLayout.xls	MaineAlt*YYYY*LOCdist.xls

IX. Non-Functional Requirements

A. Operational Requirements

1) Vendor system
   a) Performance shall be satisfactory
   b) Availability shall be uninhibited during the open windows
   c) Security measured shall be in place for protection of data and transfers
   d) Usability of the system must be satisfactory
   e) Integrity of the system shall be adequate
2) Carrier vendor timeliness
   a) Material receipt is on time
   b) Material delivery is on time
3) Training
   a) Training is performed
   b) Training is available and delivered adequately
4) Systems support, and maintenance is available
5) Schedules are adhered to (include handoff schedule to and from reporting groups)
   a) Scheduled dates are agreed to and adhered to
6) Resources
   a) Availability of personnel must be adequate and permit capacity
   b) Accessibility of systems and shall be available for processing and reporting

X. Approval

RE_ Maine PAAP