

# Assessment Standard Setting

## Proficiency Levels and Cut Score Approvals

### Introduction

The Tennessee Department of Education (department) developed the TCAP assessments to measure the current mastery of students on the Tennessee state academic standards and their progress toward college and career readiness. To complete the reporting of achievement data, it is necessary to develop cut scores that will be used to assign students to performance levels on the assessments. These cut-scores are the basis of a criterion-referenced assessment, in which student performance is judged based on the expectations determined by content experts who are educators in Tennessee. During the week of July 16-19, 2024, panels of Tennessee educators convened in Nashville to review and recommend cut scores for the assessments.

After all meetings were concluded, the department, in collaboration with psychometric experts at our assessment administration vendor (Pearson), reviewed all recommended cut scores and are presenting final recommendations for review and approval.

Today, the department is presenting cut score recommendations for performance levels for the following assessments:

- Algebra I
- Geometry
- Algebra II
- Integrated Math I
- Integrated Math II
- Integrated Math III
- Optional Grade 2 English Language Arts

For each of the indicated assessments, student performance will be reported using four performance categories, which require setting three cut scores. The department recommends the performance level cut scores shown in this report for adoption by the Tennessee State Board of Education.

### Standard Setting Process

Performance levels are used to classify and describe student performance on an assessment. In order to classify student performance into the different performance levels, the following components are generally required: 1) Policy Performance Level Descriptors, 2) Range Performance Level Descriptors, and 3) cut scores. Policy performance level descriptors provide general descriptions of what students at each performance level know and what they are able to do. The range performance level descriptors (PLDs) illustrate the performance levels in terms that are specific to a grade and subject. Cut scores represent the lowest boundary of each performance level on the scale used to score the assessments.

The process of recommending performance standards for the TCAP EOC Math and Optional Grade 2 ELA assessments is based on national best practice for standard settings. The standard setting methodology used is a modification of the well-known Angoff method (Thorndike, 1971).

### ***Policy Performance Level Definitions***

Policy Performance Level Descriptors for the TCAP assessments are shown in Table 1. The titles and descriptions of the performance levels were defined to be part of a cohesive assessment system and provide general descriptions of student performance without regard to subject or grade-level. These policy level descriptors are consistent across all grades and subject areas included in the full suite of TCAP assessments.

**Table 1. Policy Performance Level Descriptors (general education)**

<b>Performance Level</b>	<b>Policy Performance Level Descriptors</b>
<b>Level 4: Exceeds Expectations</b>	Performance at this level demonstrates that the student has an <b>extensive</b> understanding and has an <b>expert</b> ability to apply the grade/course-level knowledge and skills defined by the Tennessee Academic Standards.
<b>Level 3: Meets Expectations</b>	Performance at this level demonstrates that the student has a <b>comprehensive</b> understanding and has a <b>thorough</b> ability to apply the grade/course-level knowledge and skills defined by the Tennessee Academic Standards.
<b>Level 2: Approaching</b>	Performance at this level demonstrates that the student has an <b>approaching</b> understanding and has a <b>partial</b> ability to apply the grade/course-level knowledge and skills defined by the Tennessee Academic Standards.
<b>Level 1: Below</b>	Performance at this level demonstrates that the student has a <b>minimal</b> understanding and has a <b>nominal</b> ability to apply the grade/course-level knowledge and skills defined by the Tennessee Academic Standards.

### ***Performance Level Descriptors (PLDs)***

A multi-step iterative process was used in developing, reviewing, and approving the subject and grade-level Range PLDs. Prior to the standard setting committee, a draft set of PLDs representing a gradual increase in expectations across the performance levels was created by content staff from the department. The resulting draft PLDs were reviewed by Tennessee educators during a review committee. The committee reviewed for alignment to the Tennessee policy performance level descriptors, the Tennessee Academic Standards, and consistency of expectations across performance levels. Recommendations were made by the committee for edits and additions to the draft. Following applications of committee recommendations, a revised draft of the PLDs were reviewed and finalized by content staff from the department. Panelists who participated in the PLD review committees were asked to participate in the eventual standard setting committee.

### ***Standard Setting Meetings<sup>1</sup>***

From July 16 through July 19, 2024 following the spring 2024 operational administration, a standard setting committee meeting was conducted to provide cut score recommendations for the TCAP assessments for EOC math and optional Grade 2 ELA. There were 7 committees, with each recommending cut scores for one assessment. Each committee was composed of between 8 and 13 individuals, including teachers and non-teacher Tennessee educators (e.g., administrators, curriculum specialists, postsecondary faculty), for a total

<sup>1</sup> Full standard setting report is available for your review upon request.

of 77 educators across all eight CORE regions of the state. The participants were selected for the standard setting committee to provide content and grade-level expertise during the committee meeting and be representative of the state teaching population, including geographic region, gender, ethnicity, educational experience, community size, and community socio-economic status. Many of the panelists who worked on the initial EOC math Standards Revision Committee were recruited to participate in both the PLD Review and Standard Setting meetings for those courses. Additionally, the majority of panelists who participated in the EOC math Standard Setting meetings for the traditional courses also participated in the Standard Setting meetings for the Integrated Courses to help ensure consistency of review in light of common standards throughout the courses in each pathway.

The Extended Modified (Yes/No) Angoff standard setting method was used for the standard setting meeting (Davis & Moyer, 2015; Plake, Ferdous, Impara, & Buckendahl, 2005). This is a content- and question-based method that leads participants through a standardized process through which they consider student expectations, as defined by PLDs, and the individual questions that could be administered to students to recommend cut scores for each performance level. The standardized process was used by the committees for each grade/subject.

The process started with participants reviewing the design of the TCAP assessments and assessment questions. Panelists then created borderline descriptions based on the performance level descriptions specific to the subject and grade level to create descriptors of the knowledge, skills, and abilities that students with performance at the borderline of the performance level would be expected to demonstrate. Panelists then began an iterative judgment process where they completed three rounds of judgments on each question included on the test. The cut score recommendation for each individual participant was the expected raw score a student with performance at the borderline of the respective performance level would likely earn, calculated as the sum of the individual question judgments. Each recommended cut score from the standard setting committee was the median of the recommendations from the individual participants in the committee.

Final committee recommended cut scores were then reviewed by the department, resulting in a small number of adjustments for some performance levels to ensure final performance level cuts reflected the expectation of the Tennessee educator panelists for similar distributions of performance levels across grade levels. This final, internal review of the cut score recommendations was conducted by TDOE stakeholders, including content experts, Psychometricians and TDOE leadership, to ensure consistent expectations for performance levels across each grade-level assessment (vertical alignment). The adjustments to the recommendations made during the final review process were influenced by a desire to honor the content-based recommendations of the standard setting process, maintain high expectations for performance across the TCAP assessments grade-to-grade, and ensure the relationship among standards was coherent and defensible.

### ***Technical Advisory Committee Review***

The plans for each standard setting committee meeting were reviewed, discussed, edited, and approved in advance of each meeting by the department's Technical Advisory Committee (TAC). The TAC includes nationally recognized assessment and psychometric experts. The TAC has approved the approach

implemented in this standard setting process was high quality and the State Board of Education should feel confident in the panelists’ resulting recommendations for cut scores.

**Panelists**

A total of 52 educators from the state of Tennessee participated as panelists in the standard setting committee meetings in the 77 panel seats filled. As referenced previously, many of the EOC math panelists served on two committees, one each in the two pathways. The panelists were recruited for participation starting in the spring of 2024 and were selected using a comprehensive application process in which they demonstrated in-depth content knowledge of the Tennessee academic state standards. Panelists were selected, to the extent possible, so that they constituted a representative sample of the state teaching population in terms of region, gender, ethnicity, community size and socioeconomic status, and experience. The educators represented school districts across seven CORE regions of the state.

**Recommended Cut Scores**

The results from the standard setting meetings for the EOC mathematics and optional Grade 2 ELA committees, reviewed and finalized by the department, are presented in the tables below.

**Table 2: EOC Math Recommended Cut Scores**

Course	Performance Level			Max Points Available
	Level 2: Approaching	Level 3: On Track	Level 4: Mastered	
	Cut Score	Cut Score	Cut Score	
<b>Algebra I</b>	13	25	34	40
<b>Geometry</b>	10	19	31	40
<b>Algebra II</b>	10	22	35	40
<b>Integrated Math I</b>	12	24	34	40
<b>Integrated Math II</b>	10	21	32	40
<b>Integrated Math II</b>	10	19	30	40

**Table 3: Optional Grade 2 ELA Recommended Cut Scores**

Grade	Performance Level			Max Points Available
	Level 2: Approaching	Level 3: On Track	Level 4: Mastered	
	Cut Score	Cut Score	Cut Score	
<b>2</b>	12	24	34	40

### Reporting Scale

The process of determining the transformation rules from the Item Response Theory (IRT) scale to the final reporting scale will be guided by several principles, in order to ensure consistent understanding and ease of interpretation of the scale scores by students, parents, educators, and leaders across the state.

1. The final cut scores determined while selecting the final scaling solution should respect the cut score recommendations from the standard setting committee as closely as possible.
2. The scaling solution should involve a single linear transformation, from the IRT scale to the reporting scale.
3. The reporting scaled score range should be the same across grades and tests.

After the standards setting and finalization of recommendations were complete, the procedures for transforming student raw scores from the administration to the reporting scale were found. To ensure consistent interpretation and use of scale scores, it has been determined that the TCAP reporting scale will have the following properties across all grades:

### Reporting scale

- The lowest obtainable scale score (LOSS) would be set at 200.
- The highest obtainable scale score (HOSS) would be set at 450.

The reporting scale was set using the two fixed scores for the Lowest (LOSS) and Highest (HOSS) obtainable scale scores at 200 and 450 respectively. The scale score for all cut scores are set empirically. While the cut scores were defined with the same policy descriptions across the courses (for each type of assessment), they are not identical, and direct comparisons through averaging and aggregation across grades should not be made without study and/or statistical adjustments.

Tables 4 and 5 below present the results from the final scaling solutions for the EOC math and Optional grade 2 ELA tests respectively.

**Table 4: Reporting Scale Cuts for TCAP EOC Math**

Course	Performance Level		
	Level 2: Approaching	Level 3: On Track	Level 4: Mastered
Algebra I	300	330	352
Geometry	295	327	350
Algebra II	273	308	339
Integrated Math I	295	333	364
Integrated Math II	272	307	329
Integrated Math II	295	337	370

**Table 5: Reporting Scale Recommendations for Optional Grade 2 ELA**

Grade	Performance Level		
	Level 2: Approaching	Level 3: On Track	Level 4: Mastered
2	288	324	371

## Bibliography

Davis, L. L. & Moyer, E. L. (2015). PARCC performance level setting technical report. Available from Partnership for Assessment of Readiness for College and Careers (PARCC), Washington, D.C.

Plake, B. S., Ferdous, A. A., Impara, J. C., & Buckendahl, C. W. (2005). Setting Multiple Performance Standards Using the Yes/No Method: An Alternative Item Mapping Method. Meeting of the National Council on Measurement in Education. Montreal, Canada.

Thorndike, R. L. (Ed.). (1971). Educational measurement (2nd ed.). Washington, D.C.: American Council on Education.