BODY OF EVIDENCE PEER REVIEW SCORING GUIDE

ALIGNMENT

☐ Meets Criteria (ALL bullets)	Evidence in Plan That May Support Criteria	☐ Does Not Meet Criteria Highlight bullet(s) that apply	Recommendations
• There is documentation of **adequate sampling** of benchmarks as well as all the standards within the two representative content areas. • There is evidence of a **two-way alignment process**: all assessment items and tasks align to standards and are represented in the assessments within the two representative content areas. • The assessments from the representative content areas reflect the **cognitive depth** of the content standards and the types of student performance described in the performance standards.	☐ Assessment samples for the representative content areas (1 core & 1 non-core) are included. ☐ Blueprints for the assessment samples are included in the plan. ☐ Matrices indicating all the assessments in the representative content areas (1 core & 1 non-core) and the standards and benchmarks assessed by each are included. ☐ The processes used by the district to ensure alignment of current standards and benchmarks as well as future changes are described. ☐ If the district Body of Evidence system includes course-based information (e.g., grades), the process for assuring alignment among the course curriculum, standards, assessments, and grading practices are described and appropriate polices included. ☐ Evidence of the processes used to ensure alignment of assessment items/tasks to the levels of cognition called for in the performance standards is present. ☐ Evidence of "think aloud" protocols and/or careful examination of student work is used to evaluate/document, and revise, if necessary, the alignment of its standards and assessment system.	• The district provides little, incomplete, unclear or no evidence of **adequate sampling**. • The district provides little, incomplete, unclear or no evidence of **two-way alignment.** • The district provides little, incomplete, unclear or no evidence that the assessments reflect the **cognitive depth** of the content standards and the types of student performance described in the performance standards.	

CONSISTENCY

☐ Meets Criteria	Evidence in Plan That May Support Criteria	☐ Does Not Meet Criteria (Highlight bullets that apply)	Recommendations
• For **open-ended assessments**, the district plan describes clear procedures to be used to ensure inter-rater reliability and defines a desired, acceptable rate. Data are presented that support implementation of the stated procedures. • For **closed-ended assessments,** the district plan describes clear procedures to be used to ensure reliability and defines a desired, acceptable rate. Data are presented that support implementation of the stated procedures. • If **teacher judgment** is part of the plan**,** the plan describes procedures to ensure reliability of judgment across assessments within a course & across teachers. There is clear documentation that judgment is anchored to the performance standards. Data are presented that support implementation of the stated procedures.	☐ The procedures used to ensure inter-rater reliability on open-ended assessments are described. ☐ Inter-rater reliability data that meets acceptable rates (inter-rater reliability to meet or exceed 80% exact agreement and 98% exact + adjacent agreement) is included. ☐ The procedures used to ensure reliability on closed-ended assessments are described. ☐ Desired, acceptable rates of reliability on closed-ended assessments are stated. ☐ Reliability data on closed-ended assessments (to meet or exceed average reliability coefficients greater than 0.85) is included. ☐ Procedures used to ensure reliability of teacher judgment across assessments within a course and across multiple teachers are described. ☐ Reliability data of teacher judgment is included.	• The district provides little, incomplete, unclear or no evidence of procedures to be used to ensure inter-rater reliability on **open-ended assessments.** • The district provides little, incomplete, unclear or no evidence of procedures to be used to ensure reliability on **closed-ended assessments**. • The district provides little, incomplete, unclear or no evidence of procedures to be used to ensure reliability of **teacher judgment.** • The district provides little, incomplete, unclear or no evidence of desired, **acceptable rates of reliability** being defined. • The district provides little, incomplete, unclear or no evidence of **data** that supports implementation of the stated procedures.	

FAIRNESS

☐ **Meets Criteria** (ALL bullets)	**Evidence in Plan That May Support Criteria**	☐ **Does Not Meet Criteria** (Highlight bullets that apply)	**Recommendations**
• There is evidence the district uses procedures or tools to ensure that **assessment items/tasks** are not **biased** against subgroups of students. • There is evidence the district uses **accommodations** appropriately. • There is evidence the district provides **multiple assessment opportunities**. • A **variety of assessment formats and strategies** are included in the system. • The district **disaggregates assessment results** (i.e. ethnicity, gender & socio-economic status) and the results are used to search for possible bias in the system. • Relevant district data are presented to document that **participation rates** are at least 95% for all subgroups.	☐ The procedures (e.g., bias committees) used to ensure that items and tasks are not biased against any subgroups of students are described. ☐ Sample forms and/or notes from bias review committee meetings are included. ☐ Policies and procedures for ensuring fair participation of all students in the system (e.g. students with disabilities or English language proficiency) are evident. ☐ There is evidence that illustrates accommodations are used. ☐ There is evidence that the district system provides students with multiple opportunities, using different formats and strategies, to demonstrate their knowledge and skills. ☐ The plan includes disaggregated assessment results by identifiable subgroups (i.e. ethnicity, gender & socio-economic status) and describes how the district uses the information to make decisions. ☐ There is evidence that disaggregated assessment results are used to search for potential bias in the assessment system. ☐ The plan includes participation rates data for the content area assessments submitted.	• The district provides little, incomplete, unclear or no evidence of plans, procedures, or tools to ensure that **assessment items/tasks** are not biased against any subgroups of students. • The district provides little, incomplete, unclear or no evidence that **accommodations** are used appropriately. • The district provides little, incomplete, unclear or no evidence that **multiple assessment opportunities** are provided. • The district provides little, incomplete, unclear or no evidence that a **variety of assessment formats and strategies** are included in the system. • The district provides little, incomplete, unclear or no evidence of a process being used to **disaggregate assessment results** and the results are being used to search for possible bias in the system. • The district provides little, incomplete, unclear or no evidence that **participation rates** are at least 95% for all subgroups.	

STANDARD SETTING

☐ **Meets Criteria** (ALL bullets)	**Evidence in Plan That May Support Criteria**	☐ **Does Not Meet Criteria** (Highlight bullets that apply)	**Recommendations**
• The district plan describes a **rationale** and a **defensible method** of standard-setting. It explains how the determination is made regarding proficiency levels in each content area. • The plan identifies **cut scores** for each level of performance and the method used to determine these cut scores. It shows that they are clearly tied to performance standards. • The district plan presents a timeline showing **adequate notification** to students on progress toward proficiency in each content area. • There is evidence that the district has included **key stakeholders** (e.g., parents, community members, teachers) in the standard-setting process.	☐ The rationale and the standard-setting method used for determining proficiency at the content level is described. ☐ The cut scores used for each level of proficiency in the representative content area are included in the plan. ☐ The levels at which the cut scores have been set are clearly tied to the performance descriptors for the representative content areas. ☐ How and when individual scores are aggregated to make "graduate/not graduate" decisions are explained. ☐ The plan includes the timeline the district uses for their student notification process. ☐ The plan describes how key stakeholders are involved in the standard-setting process.	• The district provides little, incomplete, unclear or no evidence of a **rationale** and a **defensible method** of standard-setting which describes how the determination of level of proficiency is made at the content level. • The district plan provides little, incomplete, unclear or no evidence of the **cut scores** used in each content area. • The district plan provides little, incomplete, unclear or no evidence that cut scores are clearly tied to performance standards. • The district plan provides little, incomplete, unclear or no evidence of a timeline showing **adequate notification** to students on progress toward proficiency in each content area. • The district plan provides little, incomplete, unclear or no evidence that **key stakeholders** have been involved in the standard-setting process.	

COMPARABILITY

☐ **Meets Criteria** (ALL bullets)	**Evidence in Plan that May Support Criteria**	☐ **Does Not Meet Criteria** (Highlight bullets that apply)	**Recommendations**
• The district provides evidence that specific procedures are in place for ensuring comparability of assessments for all students **in a given year**, regardless of classroom, program, or school in the district. • The district provides evidence that specific procedures are in place for ensuring comparability **across years.** • The district provides evidence that specific procedures are in place for **replacing assessment tasks/items** with comparable tasks/items in terms of content, focus, and cognitive demand.	☐ There is documentation of on-going district-wide trainings, common rubrics, the use of "seeded" papers, and common administration guidelines used to ensure comparability. ☐ The district has a process for ensuring the assessments are administered similarly from year-to-year. ☐ There is evidence that the district ensures that assessments are scored the same as in previous years (e.g., the use of anchor papers and common scoring rubrics, and scoring workshops for new teachers). ☐ The plan includes evidence of procedures for replacing assessment tasks/items such as the use of assessment blueprints and protocols.	• The district provides little, incomplete, unclear or no evidence that specific procedures are in place for ensuring comparability of assessments for all students **in a given year**. • The district provides little, incomplete, unclear or no evidence that specific procedures are in place for ensuring comparability **across years.** • The district provides little, incomplete, unclear or no evidence that specific procedures are in place for **replacing assessment tasks/items** with comparable tasks/items in terms of content, focus, and cognitive demand.	