States have demonstrated their leadership and commitment to ensuring the success of all students by adopting college- and career-readiness standards. To realize the potential of these standards, states require assessments that match the depth, breadth, and rigor of the standards; accurately measure student progress toward college and career readiness; and provide valid data to inform teaching and learning.

**Assessment of College and Career Readiness.** States have taken different approaches to establishing college- and career-readiness standards and to putting in place high-quality aligned assessments. Many states have adopted the Common Core State Standards (CCSS); some have modified the CCSS to meet their state’s context and needs; and others have developed standards independent of the CCSS. To provide assessments that are aligned to these standards, many states are working together through assessment consortia, while others are taking alternative paths for transition. This document is grounded in best practices for assessment development and in the research that defines college and career readiness for English Language Arts (ELA)/literacy and mathematics. Thus, regardless of each state’s approach, this document is intended to be a useful resource for any state procuring and/or evaluating assessments aligned to their college- and career-readiness standards.

**Assessment Criteria for States to Consider.** This document provides criteria for states to consider as they develop procurements and evaluate options for high-quality state summative assessments aligned to college- and career-readiness standards. The criteria build on the states’ high-quality summative assessment principles (CCSSO, 2013) which articulate their commitment to high-quality assessments aligned to college and career readiness. To assist states in operationalizing their commitment, this document pays particular attention to not only the criteria states could ask vendors to meet, but also to the evidence states could ask vendors to provide to demonstrate criteria have been – or will be – met. States will, of course, adapt these criteria to reflect their context, standards, and procurement regulations.

**Contents of this Document.** This document begins with an overview of the assessment criteria and continues with a chart containing detailed criteria and sample evidence. These criteria do not cover every area that a state would have to address in a procurement or evaluation process. Instead, they focus on the critical characteristics that should be met by high-quality assessments aligned to college- and career-readiness standards. A more comprehensive source for the development and validation of assessments is the Standards for Educational and Psychological Testing (AERA, APA, and NCME, 1999). The assessment criteria and evidence discussed herein were developed by referencing the Standards for Educational and Psychological Testing and several other key sources listed in the bibliography. Additional state-specific criteria at the end of the document highlight a few of the most important additional issues that states may wish to consider in a procurement or evaluation process.

**Notes about Evidence and Terminology.** This document is intended to support states in selecting assessments that meet a high bar for quality. Thus, the document suggests the evidence that states will need to review in order to make informed judgments on vendors’ claims about the quality of their proposed assessments. Of course, vendors may propose assessments that are yet to be developed, assessments in development, and/or existing assessments. In designing procurement or evaluation procedures, states may therefore find it helpful to design the process for awarding “points” so as neither to reward existing (but poor quality) tests just because they have data available, nor to reward well-intentioned conceptual designs that are not executable. To support this goal, vendors should be asked to provide the most rigorous level of evidence they have available, consistent with the stage of assessment development they are in. The types of evidence that vendors should be expected to provide at different stages of development are described below:
• For assessments to be newly created, the most rigorous level of evidence will include the vendor’s descriptions of their established and proven processes; data from similar assessments; proposed test blueprints and other specifications (e.g., test design documents, test specifications, item specifications, scoring specifications); exemplar test items, passages, and forms; proposed studies, reports, and technical documentation to be created during assessment development and operation; and the processes for responding to such data. In addition, the vendor’s prior experience, expertise, and letters of recommendation should be included.

• For assessments that are currently in development, the most rigorous level of evidence will depend on the stage of assessment development. Evidence should include test blueprints and other specifications (e.g., test design documents, test specifications, item specifications, scoring specifications), and exemplar test items, passages, and forms. In addition, evidence should include as much of the data described below regarding pre-existing assessments as is available. Where such evidence is not available, vendors should provide descriptions of their established and proven processes; data from similar assessments, proposed studies, reports, and technical documentation to be created during assessment development and operation; and the process for responding to such data. In addition, the vendor’s prior experience, expertise, and letters of recommendation should be included.

• For pre-existing assessments, the most rigorous level of evidence will include comprehensive validity evidence; test blueprints and other specifications (e.g., test design documents, test specifications, item specifications, scoring specifications); annual technical reports; results of studies on scaling, equating, and reporting; and exemplar test items, passages, and forms.

Additionally, regardless of the stage of test development, states may find it helpful to put in place best practice quality assurance and other processes so that states can monitor quality throughout development and administration, and periodically evaluate evidence to ensure criteria are being met.

Finally, a note about terminology. In this document, the term “assessments” generally refers to the entire suite of summative assessments a state would procure – that is, tests of ELA/literacy and mathematics in each grade assessed. In sections specifically about ELA/literacy or mathematics, however, the term refers to the set of summative assessments in that content area. The terms “assessment” and “test” are often used interchangeably when discussing a single grade level/content area. Throughout the document, the term “tasks” refers to extended-response, open-ended test items; “test items” refers to the stimuli used to elicit a response through, for example, multiple-choice or constructed-response items as well as tasks; and “forms” are systematic collections of test items and tasks that comprise the testing experience for a particular student in a grade/content area.
Overview of Assessment Criteria

A. Meet Overall Assessment Goals and Ensure Technical Quality
   A.1 Indicating progress toward college and career readiness
   A.2 Ensuring that assessments are valid for required and intended purposes
   A.3 Ensuring that assessments are reliable
   A.4 Ensuring that assessments are designed and implemented to yield valid and consistent test score interpretations within and across years
   A.5 Providing accessibility to all students, including English learners and students with disabilities
   A.6 Ensuring transparency of test design and expectations
   A.7 Meeting all requirements for data privacy and ownership

B. Align to Standards – English Language Arts/Literacy
   B.1 Assessing student reading and writing achievement in both ELA and literacy
   B.2 Focusing on complexity of texts
   B.3 Requiring students to read closely and use evidence from texts
   B.4 Requiring a range of cognitive demand
   B.5 Assessing writing
   B.6 Emphasizing vocabulary and language skills
   B.7 Assessing research and inquiry
   B.8 Assessing speaking and listening
   B.9 Ensuring high-quality items and a variety of item types

C. Align to Standards – Mathematics
   C.1 Focusing strongly on the content most needed for success in later mathematics
   C.2 Assessing a balance of concepts, procedures, and applications
   C.3 Connecting practice to content
   C.4 Requiring a range of cognitive demand
   C.5 Ensuring high-quality items and a variety of item types

D. Yield Valuable Reports on Student Progress and Performance
   D.1 Focusing on student achievement and progress to readiness
   D.2 Providing timely data that inform instruction

E. Adhere to Best Practices in Test Administration
   E.1 Maintaining necessary standardization and ensuring test security

F. State Specific Criteria (as desired)
   Sample criteria might include
   • Requiring involvement of the state’s K-12 educators and institutions of higher education
   • Procuring a system of aligned assessments, including diagnostic and interim assessments
   • Ensuring interoperability of computer-administered items