

## CAEP Standard 3, Component 3.2 measures of academic proficiency

June 2017

This paper contains CAEP’s official publication of performance criteria for approved nationally-normed or other substantially equivalent academic achievement measures used for group average performance evidence to document Standard 3, Component 3.2. An additional document, [Guidelines for Equivalence Studies Conducted for CAEP Standard 3, Component 3.2 Academic Proficiency Measures](#), contains guidance for sponsors of studies that are intended to demonstrate substantial equivalence of additional assessments, together with templates for CAEP reviewers of those studies. Both were prepared at CAEP’s request by Dr. William Lorié, psychometric consultant and research scientist, to assist educator preparation providers (EPPs) as they conduct self-study reports.

### CAEP’s published measures for use in Component 3.2

Accreditation Standard 3 of the Council for the Accreditation of Educator Preparation (CAEP) enables education preparation providers (EPPs) to demonstrate that “the quality of candidates is a continuing and purposeful part of its responsibility from recruitment, at admission, through the progression of courses and clinical experiences, and to decisions that completers are prepared to teach effectively and are recommended for certification” (CAEP, 2013).

Component 3.2 of Standard 3 (amended, 2016) specifies minimum criteria for academic achievement which must be met by the group average performance of candidates whose preparation begins during an academic year in each accredited EPP. The academic proficiency criterion includes an assessment portion, whereby each EPP demonstrates that their candidates’ average performance is at or above the 50<sup>th</sup> percentile in mathematical, reading, and (beginning in 2021) writing achievement.

CAEP publishes the Standard 3, Component 3.2 test score requirements for the academic group average, with currently approved tests. These requirements are updated each year.

The most recent version (2016) of the academic proficiency criterion was drafted in a way that allows EPPs considerable flexibility in assessments and the point in time when they are applied (i.e., from prior to application to just before completion). In addition to the SAT and ACT – tests described as “examples” of nationally normed assessments that EPPs might use for Component 3.2 in the original (2013) version of that Component – EPPs may now use substantially equivalent alternative measures of mathematical, reading and writing achievement. In the interest of fairness and uniform applicability of Component 3.2 across EPPs, CAEP has established a process by which assessments can be reviewed for substantial equivalence.

Once a test has been approved through this process, test scores for this assessment will be published together with all other recognized Standard 3, Component 3.2, test scores. EPPs may use any of the assessments listed here to demonstrate that they have met the requirements of Standard 3, Component

3.2. Note that the requirements must be met for each domain separately – reading, mathematics, and writing.

**Chart 1: Requirements for Average Scores on Assessments Approved by CAEP for Demonstrating Standard 3, Component 3.2**

Assessment	Test or Section	3.2 Domain—NOTE: proficiency must be met for each domain	Group average performance requirements of candidates whose preparation began during the 2016-2017 academic year or earlier
ACT	“Reading”	Reading	21.25
ACT	“Math”	Math	21.25
ACT	“Writing”	Writing	6.60
SAT	“Evidence-Based Reading and Writing”	Reading	543.33
SAT	“Math”	Math	532.50
SAT	“Essay – Writing dimension”	Writing	5.30
GRE	“Verbal Reasoning”	Reading	150.75*
GRE	“Quantitative Reasoning”	Math	152.75*
GRE	“Analytical Writing”	Writing	3.74*

\* GRE average score requirements are based on GRE user norms and should not be interpreted as equivalent to the average score requirements of other assessments due to differences in the population of test takers. The test user reference group has completed requirements for a baccalaureate degree and has, in general, four additional years of academic preparation compared with the students typically taking college entrance tests. See final paragraphs in the section below about reference groups.

### Terminology and Concepts in Academic Achievement

The *assessment portion of the academic proficiency criterion* (or simply, the *criterion*) is “a group average performance on nationally normed assessments or substantially equivalent state-normed assessments of mathematical, reading, and writing achievement in the top 50 percent of those assessed.”

There are four aspects of the criterion: A construct encompassing three distinct domains, an instrument description, a performance standard, and a reference group.

- The *construct* is “mathematical, reading and (beginning in 2021) writing achievement,” which is achievement in three distinct *domains*. EPPs must meet the criterion for each domain separately.
- The *instrument description* is “nationally normed assessments ... of mathematical, reading and writing achievement” and is qualified by “substantially equivalent state-normed assessments of mathematical, reading, and writing achievement.”

- The *performance standard* in the criterion is “group average performance ... in the top 50 percent of those assessed.”
- The *reference group* is “those assessed.” In the context of the previous (2013) version of the criterion, and after discussions / research as part of a consulting engagement with W. Lorie, CAEP determined that the reference group for the criterion is the national population of college-bound 11<sup>th</sup> and 12<sup>th</sup> grade students<sup>1</sup>. For reading and math, CAEP has determined that this college-bound group is satisfactorily represented, as a proxy measure, by a population that the College Board defines as its national “user percentile.” More specifically, the College Board’s population definition is “U.S college-bound students in the 11<sup>th</sup> and 12<sup>th</sup> grades, weighted to represent students who typically take the SAT last as 11<sup>th</sup> or 12<sup>th</sup> graders.” For writing, the current CAEP proxy measure is ACT test-takers in 2016, a group made up of nearly two-thirds of all high school graduates.

Selection of a reference group directly aligned with CAEP’s definition (the final aspect of the Component 3.2 criterion, above) is challenging. The high school students taking ACT tests represent a wider segment of the high school population than CAEP’s college-bound definition. That has the effect of lowering the 50<sup>th</sup> percentile score compared with the college-bound group<sup>2</sup>. Also, SAT data are reported for a user group that serves as a satisfactory proxy for CAEP’s. Yet CAEP acknowledges that increasing adoption of the SAT by whole states, as a requirement, may reduce the extent to which a future SAT College Board report for the college-bound group resembles the CAEP reference group. **CAEP ensures against this type of change in the proxy used for college entrance tests by setting the 2016 cut scores as the lower bound for any future adjustments to the 50<sup>th</sup> percentile scores.**

Note, however, that this college entrance test rule does not apply to use of GRE for initial teacher preparation at the graduate level. The reference group for GRE represents the obverse situation. The

---

<sup>1</sup> The term “college-bound” refers to 11<sup>th</sup> and 12<sup>th</sup> graders who take college entrance tests because they are thinking about applying, or intending to apply, to attend college. College-bound should not be interpreted as “admitted” or expecting to enroll—a narrower group. Also, at the opposite extreme, it does not mean all 11<sup>th</sup> and 12<sup>th</sup> graders, e.g., in those states that use college entrance tests as part of their student progress and attainment measures. While EPP candidates are not limited to those who come directly from high school, the CAEP reference population is directly associated with the primary population preparing to teach, and for whom the college experience will be their gateway to teaching. The college entrance exams that many of these applicants take also serve as anchors for comparisons with many other tests, enhancing the range of opportunities to document “substantial equivalence” under the Component 3.2 requirement.

<sup>2</sup> In its annual publication of ACT results for high school students, *the Condition of College & Career Readiness, 2016 National ACT* (URL: [http://www.act.org/content/dam/act/unsecured/documents/CCCR\\_National\\_2016.pdf](http://www.act.org/content/dam/act/unsecured/documents/CCCR_National_2016.pdf)) ACT provides the follow description: “Nearly two-thirds (64%) of all 2016 US high school graduates took the ACT, up from 59% last year and from 40% in 2006. As a result, this year’s data are based on a more representative sample of the nation’s graduating class than ever before.

- Overall achievement levels—both the average ACT “Composite” score and the percentages of students meeting the ACT College Readiness Benchmarks—dropped this year compared to last year, likely due to the significant increase in the percentage of students tested. The addition of seven more states that required all students to take the ACT in the 2016 data set means students from a broader range of academic preparation were tested, including many students who may not have been preparing to attend college. This decline in scores nationally does not necessarily represent an actual change in national performance or achievement, but rather reflects the changes in the makeup of the testing population.”

test taker population is more selective, comprised of individuals who have completed or are nearing completion of a baccalaureate degree, who have experienced four additional years of academic preparation compared with the SAT and ACT test takers. CAEP will continue to look for refinements in test taker population reference groups but believes the ones used in this paper are responsibly described, transparent, and practical for EPPs to use.

### **Requirements for Assessments to be Deemed “Substantially Equivalent”**

The academic proficiency criterion contained in CAEP’s Component 3.2 specifies use of “nationally-normed assessments or substantially equivalent state-normed assessments of mathematical, reading, and (beginning in 2021) writing achievement.” The Component goes on to note that “CAEP will work with states and providers to designate, and will periodically publish, appropriate ‘top 50 percent’ proficiency scores on a range of nationally or state normed assessments and other substantially equivalent academic achievement measures.” To implement these provisions, CAEP has adopted guidelines and procedures for a *substantial equivalence review* that states, testing organizations and other potential sponsors can follow to substantiate their claim that tests would be appropriate for use as Component 3.2 evidence.

There are four requirements for tests to be deemed substantially equivalent for the purposes of Standard 3, Component 3.2. These must be met separately for mathematical, reading, and/or writing achievement (the three domains). If an instrument assesses more than one of these domains, it is possible for it to be substantially equivalent for just one domain, or two, or all three.

**Requirement 1.** *The content, format, and depth of knowledge of the assessment is (A) comparable to that of any from a set of conforming frameworks for the domain, which are considered by CAEP as sufficiently-vetted operational definitions of frameworks for achievement for that domain; OR (B) appropriate as a measure of the domain, and strongly supported as such by validity argument and supporting research.*

EPPs can use the New SAT “Evidence-Based Reading and Writing” section scores and the ACT “Reading” test scores as evidence of meeting the criterion for reading achievement. Similarly, EPPs can use the New SAT “Math” section scores and the ACT “Math” test score as evidence of meeting the criterion for mathematical achievement. For writing, EPPs can use the ACT “Writing” test and the SAT “Essay” as evidence of meeting the criterion for writing achievement. These six tests are *conforming assessments*. They meet several requirements to be conforming assessments, one of which is that they are built to *conforming content frameworks*. This means that the content, format, and depth of knowledge (DOK) of the assessment producing the score is considered by CAEP as a sufficiently vetted operational definition of a suitable framework for the criterion, for the domain of the assessment.

**Conforming frameworks are listed in Chart 2. These conforming assessments also serve as base points for sponsors of studies seeking to document “substantial equivalence.”** Note that not all assessments referenced in the conforming framework list are eligible assessments for Component 3.2. This is because some other requirements have not (yet) been met for these assessments. For example, an assessment might not provide an appropriate 50<sup>th</sup> percentile of the reference group of Standard 3, Component 3.2. If an assessment of reading is on the list of conforming frameworks for reading, this means that the assessment tests reading in such a way that CAEP considers it an appropriate measure for reading achievement for the target population of Standard 3, Component 3.2.

The purpose of this requirement is to ensure that alternative assessments proposed are defensible as measures of the domains referenced in Standard 3, Component 3.2.

To meet the requirement an instrument sponsor should submit documentation supporting content validity claims – i.e., that the instrument assesses the target domain in an appropriate and defensible manner. The instruments in CAEP’s conforming frameworks lists serve as benchmarks and examples. To the extent that the content, format, and depth of knowledge of a proposed assessment departs from all listed conforming frameworks, more independent construct and content validation support is needed to support a positive substantial equivalence outcome.

**Requirement 2.** *The instrument must be sufficiently reliable, fair, and free from bias for any subgroup in the population.*

Reliabilities of section scores of the SAT, test scores of the ACT (not the composite), and content area scores of PARCC or Smarter Balanced high-school assessments can be considered benchmarks for acceptable reliability under Standard 3, Component 3.2. Evidence that assessments are fair and unbiased can consist of summaries from differential item functioning (DIF) statistics and findings from data review meetings in which items flagged for DIF are discussed.

**Requirement 3.** *Score percentiles are available for the reference group of Standard 3, Component 3.2, national college-bound 11<sup>th</sup> and 12<sup>th</sup> graders, or for an acceptable proxy measure. For reading and math the accepted proxy measure is the SAT population reported by the College Board as the “national user percentile” for 2016; for writing the current proxy is ACT test-takers in 2016. Moreover, a score on the test being reviewed has been proposed as meeting the group average 50<sup>th</sup> percentile referenced in Standard 3, Component 3.2. If the 50<sup>th</sup> percentile is determined by linking or concordance to another assessment, then an equivalence table should be provided.*

This third requirement can be met either directly by the instrument – for example if an assessment program has systematically sampled this reference group and has conducted a norming study – or indirectly, by some sort of linking or concordance. The linking should be to scores on a scale that reports on the reference group—national college-bound 11<sup>th</sup> and 12<sup>th</sup> graders; is an acceptable proxy of the reference group (*which for reading and math is the “national user percentile” for 2016 by the College Board, and for writing is ACT test-takers in 2016*); or a scale that is linked to an acceptable proxy.

**Requirement 4.** *In all other respects, the assessment should be defensible per the latest edition of the [Standards for Educational and Psychological Testing](#).*

**Chart 2. Conforming Frameworks** [NOTE: See section above labeled “Requirement 1”. The assessments listed in this table are considered by CAEP as sufficiently vetted operational definitions of a suitable framework [i.e., in terms of the content, format, and depth of knowledge (DOK)] for the specified assessment domain. These assessments also serve as base points for sponsors of studies that would document “substantial equivalence.” Those that appear in Chart 1, above, are eligible assessments for Component 3.2 evidence, and the appropriate Component 3.2 50<sup>th</sup> percentile score is specified.

Assessment Program	Reading Achievement	Mathematical Achievement	Writing Achievement	Reference Documents
New SAT (March 2016 or later)	“Evidence-Based Reading and Writing” (Section score)	“Math” (Section score)	“Essay” (Reading, analysis and Writing scores)	The College Board (2015). <i>Test Specifications for the Redesigned SAT</i> . <a href="https://collegereadiness.collegeboard.org/pdf/test-specifications-redesigned-sat-1.pdf">https://collegereadiness.collegeboard.org/pdf/test-specifications-redesigned-sat-1.pdf</a> The College Board, SAT “Essay”, retrieved at: <a href="https://collegereadiness.collegeboard.org/sat/inside-the-test/essay">https://collegereadiness.collegeboard.org/sat/inside-the-test/essay</a> The College Board, SAT “Essay” Scoring, retrieved at: <a href="https://collegereadiness.collegeboard.org/sat/scores/understanding-scores/essay">https://collegereadiness.collegeboard.org/sat/scores/understanding-scores/essay</a>
ACT	“Reading” (Test score)	“Math” (Test score)	“Writing” (Test score)	ACT (2014). <i>Technical manual: The ACT</i> . Chapters 1 and 2. <a href="https://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf">https://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf</a>
Smarter Balanced	“English Language Arts / Literacy” (Test score)	“Mathematics” (Test score)	Not currently identified	Smarter Balanced Assessment Consortium (2015). <i>Content Specifications for the Summative Assessment of the Common Core State Standards for English Language Arts and Literacy in History/Social Studies, Science, and Technical Subjects</i> . <a href="https://www.smarterbalanced.org/wp-content/uploads/2015/08/ELA_Content_Specs.pdf">https://www.smarterbalanced.org/wp-content/uploads/2015/08/ELA_Content_Specs.pdf</a>  Smarter Balanced Assessment Consortium (2015). <i>Content Specifications for the Summative Assessment of the Common Core State Standards for Mathematics</i> . <a href="https://www.smarterbalanced.org/wp-content/uploads/2015/08/Mathematics-Content-Specifications.pdf">https://www.smarterbalanced.org/wp-content/uploads/2015/08/Mathematics-Content-Specifications.pdf</a>
PARCC	“English Language Arts / Literacy” (Test score)  “Reading” (Test score)	“Mathematics” (Test score)	“Writing” (Test score)	PARCC (2016). “ELA Test Specifications Documents” [web page]. PARCC Grades 6-11 High Level Blueprints, PARCC Grade 11 Reading Evidence Tables, and PARCC Grade 11 Writing Evidence Tables. <a href="http://www.parcconline.org/assessments/test-design/ela-literacy/test-specifications-documents">http://www.parcconline.org/assessments/test-design/ela-literacy/test-specifications-documents</a>  PARCC (2016). “Math Test Specifications Documents” [web page]. PARCC Mathematics High Level Blueprint; Mathematics Claim Structure Documents: High School; PARCC Mathematics Evidence Statements for the Algebra I – Geometry – Algebra II path or the Mathematics I-II-III path; and PARCC Informational Guides for either of the paths. <a href="http://www.parcconline.org/assessments/test-design/mathematics/math-test-specifications-documents">http://www.parcconline.org/assessments/test-design/mathematics/math-test-specifications-documents</a>
GRE	“Verbal Reasoning” (Section score)	“Quantitative Reasoning” (Section score)	“Analytical Writing” (Section score)	ETS (2017). “Test Content and Structure” [web page]. All links on follow-up links. <a href="https://www.ets.org/gre/revised_general/about/content/">https://www.ets.org/gre/revised_general/about/content/</a>

Note: The content, format, and depth of knowledge (DOK) (collectively called *framework*) of the tests or test sections generating the scores in this table are considered by CAEP as sufficiently vetted operational definitions of conforming frameworks for mathematical, reading, or writing achievement applied to Component 3.