Results from AAC&U's VALUE Rubrics can be used to report student learning outcomes on the VSA's College Portrait. Institutions can administer the VALUE Rubrics for VSA reporting in one of two ways:

- Value-added: a statistically representative random sample of freshmen student artifacts from their first semester is evaluated using the VALUE Rubrics and statistically representative random sample of senior student artifacts from their last semester is evaluated using the rubrics.
- Benchmark: a statistically representative random sample of senior student artifacts from their last semester is evaluated using the rubrics.

Institutions should administer the VALUE Rubrics in keeping with the needs of their local campus and in accordance with AAC&U guidelines and recommendations. Specific expectations for reporting on the VSA are outlined below, but these in no way should overrule requirements provided by AAC&U.

Demonstration Project Administration Guidelines

The AAC&U VALUE Rubrics offer institutions more flexibility in their administration than the three standardized assessment instruments approved for Student Learning Outcomes reporting on the College Portrait. The VSA recognizes both the need to keep as much flexibility for institutions as possible and the need to provide information to the public and consumer audience of the College Portraits that is comparable across institutions. With that in mind, the VALUE Rubrics Demonstration Project will include additional opportunities for dialogue across institutions and with the VSA staff for sharing of best practices and adjustments to the reporting model. Initial general guidelines are outlined below, but we encourage institutions using the VALUE Rubrics to contact us if they are using them in other ways.

Preparing the VALUE Rubrics

Results from two of the VALUE Rubrics are required for reporting on the College Portrait: Written Communication and Critical Thinking.

- The language in the Rubrics can be tailored for disciplines, but the Category (rows) and Descriptor (columns) titles need to stay the same. If you are interested in examples of tailored rubrics, please contact the VSA Support Team at support@collegeportraits.org.
- Institutions can add additional Descriptors above the Capstone level, but when reported on the College Portrait the scores should be reported as if they were given at the Capstone level (e.g., a 5 is reported as a 4).
- Institutions can add a "zero" score to represent missing data; 0 scores will be displayed on the College Portrait and labeled as missing data.
- Institutions can add Categories (e.g., to cover additional aspects of an assignment or other outcomes to be evaluated), but additional Categories will not be reported on the College Portrait.

![College Portrait of Undergraduate Education - Voluntary System of Accountability]

**Voluntary System of Accountability (VSA)
Administration and Reporting Guidelines:
AAC&U VALUE Rubrics – DEMONSTRATION PROJECT**

- Institutions **cannot** break apart existing Categories for reporting on the College Portrait; we strongly advise that institutions do not break apart and try to recombine Categories.

Raters and Ratings

Institutions should determine who will be trained as raters for the VALUE Rubrics. Raters may be course faculty, Assessment Committee members, trained graduate teaching assistants, or others institutional leaders feel are qualified to evaluate student work. The VSA recommends that each student artifact be evaluated by two raters.

Institutions aren't required to use the rubrics for grading – in fact the VSA would encourage that institutions collect artifacts from appropriate courses/assignments and evaluate them separately (outside the context of a course grade). Rating of student artifacts does not need to occur at the time the artifact is evaluated for a course, but it could if the institution felt that was the most efficient approach. Institutions may also collect student artifacts and evaluate them later (e.g., during a workshop at the end of the semester).

The same criteria apply for students at all levels, they aren't "adjusted" for expected performance of the student (e.g., Freshmen vs. Senior); it's expected that Freshmen will score mostly 1s with maybe some 2s, which would not be appropriate to equate to a grade of D or C for a Freshmen-level class)

Calibration of the VALUE Rubrics

Calibration is critical to successful use of any rubric across multiple raters. There are many examples of calibration protocols available, but they tend to follow the same general process:

1. The rubric is distributed to and reviewed by those participating in the calibration. Questions about vocabulary or interpretation of the rubric levels are discussed with the group until all participants feel comfortable distinguishing between the different Categories and Descriptors within the rubric.
2. A sample of student work (an artifact) is distributed to the reviewers; each reviewer gets a copy of the same artifact. Each reviewer reviews the artifact and uses the rubric to evaluate it.
3. When all members of the group have finished reviewing the artifact, the members discuss their ratings.
   a. Discussion of the evaluation of the artifact as a whole occurs and the overall ratings are recorded.
   b. Discussion of each Category then begins.
      i. The ratings from each reviewer are collected and recorded.
      ii. Those who rate the artifact highest or lowest are asked to explain the reason for their rating using evidence from the artifact and the language of the rubric if possible.
      iii. Others who gave different ratings are asked to explain their ratings.

     iv.   Reviewers may change their ratings during the course of the discussion. Periodically, the group should check for consensus as reviewers alter their ratings.

     v.   When consensus is reached, the group moves on to discuss the next Category using the same approach.

  c.   The purpose of the discussion is to refine the definition of each Category and Descriptor so that each reviewer has a clear understanding of how to distinguish student work at one level or another.

4.   The review process should be repeated at least one more times using a different artifact and may be repeated many times until the raters are comfortable with their ability to consistently apply the rubric to student work.

AAC&U and the VSA periodically conduct train-the-trainer rubric calibration workshops as pre-conference workshops. The VSA is also investigating the possibility of conducting a train-the-trainer webinar for rubric calibration. Additional information will be posted to the Webinars and Presentations section of the VSA website (http://www.voluntarysystem.org) as training resources are made available.

Student Artifacts and Assignments

For reporting on the College Portrait, the VALUE Rubrics can be used with multiple assignments or individual signature assignments for each student. If multiple assignments are scored (e.g., a collection of papers produced during a single course are evaluated with the Written Communication rubric), one score per student representing their performance across all assignments evaluated should be reported. Portfolios should be assigned an overall score for the entire portfolio; each portfolio counts as 1 score/student for sample size calculations.

Institutions and departments should determine appropriate assignments for evaluation. Assignment options include, but aren't limited to:

- assignments tailored by discipline (e.g., a senior capstone assignment given in Biology plus a senior capstone assignment given in Philosophy plus etc.);
- a common single assignment (the same assignment given to all students across all sections of English 111); or
- a collection of assignments within a portfolio (e.g, a collection of assignments from all freshmen *or* senior level courses; the work evaluated should all be from *either* the freshmen year *or* the senior year, not a collection of work representing the progression of student work across years).

The type and number of assignments each institution decides to use may be collected as part of the demonstration project to help the VSA compile and share examples of how institutions are implementing the VALUE Rubrics.

Score Aggregation

The distribution of outcomes across the four proficiency levels of each rubric within each department or unit should be "aggregated" to a university-level distribution. There may be several appropriate and valid methods for aggregation, but the VSA recommends that each Category be assigned a weight (to be determined by the institution and applied to all scores for a given rubric) and then averaged to provide a total overall score. For instance, if all the Categories are weighted equally and a student's score consisted of two Category scores at the Capstone 4 Level (4 points each, for a total of 8 points), two at the Milestone 3 level (3 points each, for a total of 6 points and a cumulative total of 14 points), and one at the Milestone 2 level (2 points, for a cumulative total of 16 points), their average score would be 16 points / 5 Categories = 3.2, resulting in an overall score nearest Milestone 3. This example along with two others is also shown below in spreadsheet format:

Critical Thinking	Student 1	Student 2	Student 3
Explanation of issues	4	4	3
Evidence	2	4	3
Influence of context and assumptions	3	4	2
Student's position	4	4	2
Conclusions and related outcomes	3	1	2
Total	16	17	12
Average Score	3.2	3.4	2.4
Nearest Category	Milestone 3	Milestone 3	Milestone 2

The VSA may collect from each institution the weighting criteria they used to come up with a total rubric score for each students, again for the purposes of sharing examples of how institutions are implementing the VALUE Rubrics.

Sampling Expectations

For both value-added and benchmark administrations, the sampling objective is to construct and administer the test to a representative sample for the institution by:

- Student race/ethnicity,
- Student gender,
- Percent of students receiving Pell Grants,
- Admissions test scores, and
- Broad area of study (optional).

While the sample should be representative of these student populations, the VSA will not be comparing results within student groups at a given institution so samples do not need to be stratified by these groupings. If an individual institution wishes to compare student subpopulations, they should work with the publisher to determine an appropriate stratified sampling methodology.

For institutions that evaluate their students on a rolling cycle such that over the course of multiple years students across campus are evaluated, the sample should be calculated based on the total population across all years included in the sample, not just a single year of the cycle. For instance, if 1000 students are in each year's senior class and data for a represenative sample are collected over the course of 3 years, the total population for determining the size of a representative sample would be 3000.

Representative sample size should be determined using typical statistical methods based on the size of the population being represented. The precise formula for determining sample size for the VSA College Portraits is:

$$\text{Sample Size} = 384.16/[1+(384.16/\text{population})]$$

For convenience, the table below provides representative sample sizes for different populations of students. For institutions with graduating classes greater than 2,500 students, a sample size of 350 should be used.

Recommended Sampling Requirements for Benchmark Administrations

Population Size *(total # of graduating students)*	Sample size needed
500	217
750	254
1000	278
1250	294
1500	306
1750	315
2000	322
2250	328
2500	333

Institutions may either identify students and collect artifacts only from students included in the sample or collect artifacts from all students and select a random sample from the collected artifacts. Institutions may also report results for larger samples then required so long as the sampling objective for an institutionally representative is maintained.
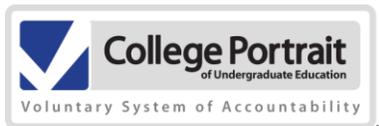
Data Required for Reporting Results

In order to provide additional context and detail about the meaning and use of student learning outcomes results on VSA participating campuses, institutions will be asked to enter the following data when reporting results on their College Portrait.

Administration Experience

Institutions will enter text to answer the following 5 questions about the results they post on their College Portrait:

1. Why did you choose the instrument you did for your institutional assessment?
2. Which students are assessed when?
3. How are assessment data identified and collected?
4. How are data aggregated/analyzed and reviewed by the institution?
5. How are assessment data used to guide program or institutional improvements?

Each answer will be displayed separately with the question as a header. The recommended character limit for each response is 750 characters. Answers should be written in language accessible to a public consumer audience and while links can be embedded, linking to long or technical reports is discouraged.

Students Tested

Institutions will enter data about the demographic make-up of both the total eligible population and the tested sample for each student group.

- The total eligible population for new entering students includes all students who could have been included in the testing population, e.g., all freshmen students who entered in the summer or fall term or all senior students expected to graduate in the spring or summer term.
- The tested sample is all students who completed the testing.

For each group, the following data will be collected and reported on the College Portrait:

- Distribution by gender,
- Distribution by race/ethnicity (using broadly defined categories of US Historically Underrepresented Minority, White/Caucasian, International, and Race/Ethnicity Unknown),
- % Low-income (as determined by Federal Pell Eligibility),
- ACT or SAT interquartile range, and
- Distribution across areas of study, as defined by the institution (optional).

AAC&U VALUE Rubric Results

Institutions can report results from both rubrics in the same year or report results from different years so long as no set of results is more than 3 years old.

Value-added results

The VSA won't compute or calculate an "expected" level of improvement, but will instead display the full distribution of scores for both freshmen and seniors on the same chart. Institutions will be provided the opportunity to discuss the differences in new student and senior student scores and how those differences fit within institutional goals.

The results will be displayed on the College Portrait approximately as shown below:
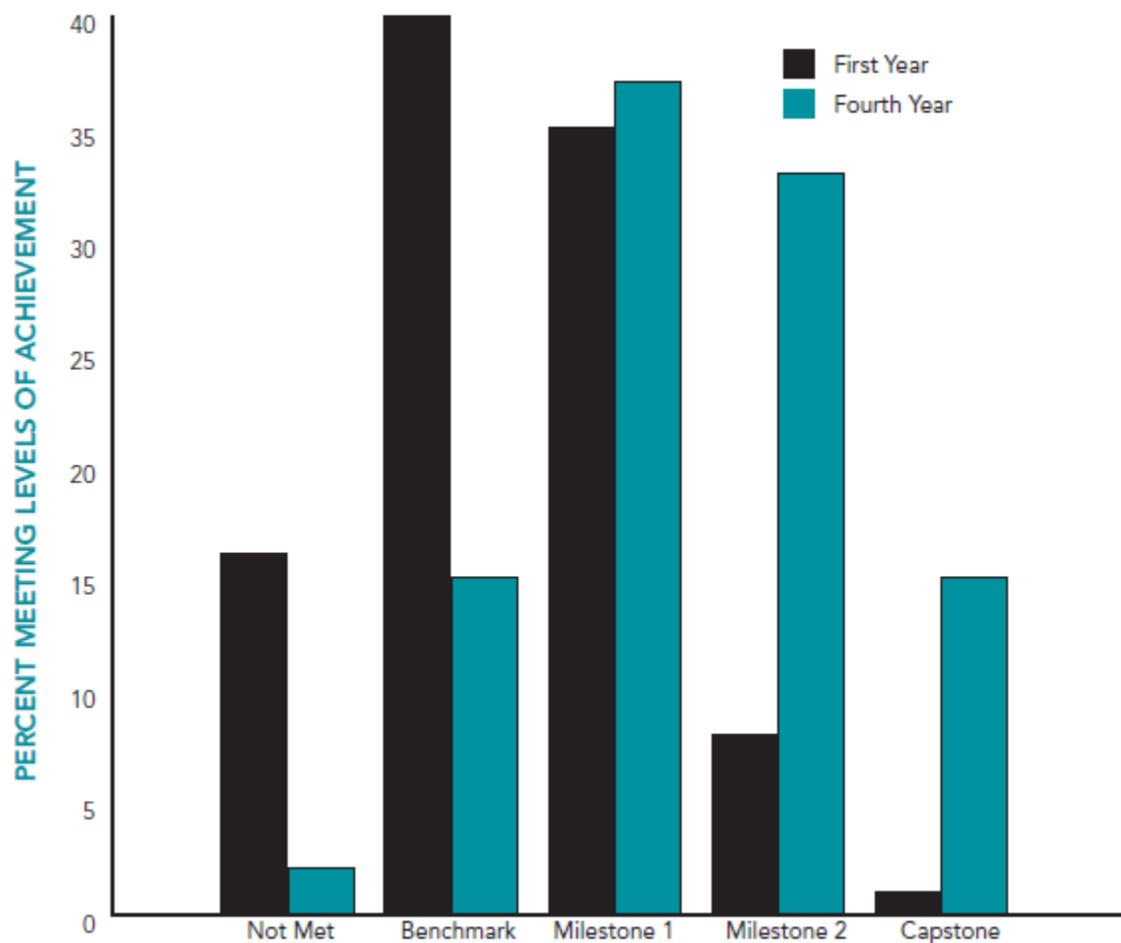
Benchmark results

Scores on the College Portrait will be displayed as the full distribution of scores for seniors. The underlying assumption is that the Capstone proficiency level represents expected senior level achievement, however, institutions will be provided the opportunity to discuss their own norms or achievement goals. The VSA encourages institutions to determine a benchmark appropriate to their

institution, e.g., 85% of graduating seniors should perform at the Capstone level, and to compare their results to their stated goal or benchmark.

Score Distribution Charts (Optional)

For institutions that wish to provide additional information about their students performance for **either** value-added or benchmark administrations, the option to provide data for and display full subscore distribution charts is available. These charts would be accessed from a link on the main results page and would consist of simple bar charts showing the full distribution of scores for the appropriate subscore.

For the AAC&U VALUE Rubrics, the following subscore charts are available:

- Critical Thinking
    - Explanation of issues
    - Evidence
    - Influence of context and assumptions
    - Student's position (perspective, thesis/hypothesis)
    - Conclusions and related outcomes (implications and consequences)
- Written Communication
    - Context of and Purpose for Writing
    - Content Development
    - Genre and Disciplinary Conventions
    - Sources and Evidence
    - Control of Syntax and Mechanics

All charts will report the distribution of student scores across the Descriptor for each subscore.