



General Education Assessment of Mathematics (Spring 2018)

History of the Assessment Program

Assessment of student learning outcomes is a national expectation in higher education, and the expectation calls for increased accountability. Section 2.7.3 of the Southern Association of Colleges and Schools' (SACS) accreditation standards requires in each undergraduate program the successful completion of a general education component that:

- 1) is a substantial component of each undergraduate degree,
- 2) ensures breadth of knowledge, and
- 3) is based on a coherent rationale.

Section 3.5.1 of the SACS accreditation standards also requires that “the institution identifies college-level competencies within the general education core and provides evidence that graduates have attained those competencies.”

Based on these standards, in 2005, the Provost charged the General Education Curriculum Committee (GECC) with developing and implementing an assessment program. To accomplish this directive, the committee developed and modified rubrics to measure student performance in the competencies stated in the preamble of the General Education Plan: “The General Education Program at the University of Louisville fosters active learning by asking students to:

- 1) think critically,
- 2) to communicate effectively, and
- 3) understand and appreciate cultural diversity.”

The GECC initiated the first General Education Assessment in fall of 2005. The university adopted LiveText© as the platform for electronic assessment of General Education artifacts in the fall of 2010. The process, results, and findings from each assessment iteration are presented to the GECC to drive continuous improvement of the university’s general education program.

Assessment Administration

The General Education Program at the University of Louisville advances three over-arching competencies: critical thinking, effective communication, and cultural diversity. In addition, the university has defined additional learning outcomes for the following content areas: Arts and Humanities, Mathematics, Natural Sciences, Oral Communication, Social and Behavioral Sciences, and Written Communication, and the Cultural Diversity competency area. The University of Louisville Student Learning Outcomes are closely aligned with the Statewide General Education Student Learning Outcomes.

The Spring 2018 assessment was focused on courses in the Mathematics content area. A crosswalk of the outcomes and assessment measures for Mathematics is provided in Appendix A

to demonstrate alignment between the assessment measures, the UofL content area outcomes, and the statewide content area outcomes.

University of Louisville Mathematics Learning Outcomes

Mathematics is concerned with solving real-world problems through mathematical methods. Students who satisfy this requirement will demonstrate that they are able to do all of the following:

1. Represent mathematical information symbolically, visually, and numerically;
2. Use arithmetic, algebraic, and geometric models to solve problems;
3. Interpret mathematical models, such as formulas, graphs, and tables;
4. Estimate and check answers to mathematical problems, determining reasonableness; alternatives; and correctness and completeness of solutions.

Statewide Quantitative Reasoning Student Learning Outcomes

1. Interpret information presented in mathematical and/or statistical forms.
2. Illustrate and communicate mathematical and/or statistical information symbolically, visually, and/or numerically.
3. Determine when computations are needed and to execute the appropriate computations.
4. Apply an appropriate model to the problem to be solved.
5. Make inferences, evaluate assumptions, and assess limitations in estimation modeling and/or statistical analyses.

University of Louisville General Education Mathematics Rubric Measures

Critical Thinking Rubric for Mathematics

- (M1) Correctly interprets mathematical information
- (M2) Applies mathematical models to solve problems
- (M3) Represents mathematical information
- (M4) Provides complete, reasonable, and correct answers

The University of Louisville General Education Rubrics use a four-point scale, with 4 indicating performance of the measure as “clearly evident,” 3 indicating performance as “usually evident,” 2 indicating “minimally evident,” and 1 indicating performance as “not evident.” In addition, a score of “not requested” could be assigned for assignments that did not provide an opportunity for the student to demonstrate the criterion within the rubric measure.

The university will transition to the Cardinal Core program in summer 2018. At that time, the Mathematics (M) content area will transition to the Quantitative Reasoning (QR) content area. New learning outcomes will be implemented at that time.

Assessment Process

For the Spring 2018 assessment of student work from the Mathematics content area, the Office of General Education Assessment collaborated with department chairs regarding the details of the upcoming assessment to ensure faculty participation and appropriate sampling. A formal memo outlining the project and process was also provided to each of the department chairs and to all faculty teaching General Education courses within the Mathematics content area prior to the start of the semester to ensure a mutual understanding of project expectations. The initial communication provided a timeline for collection of assignment prompts, answer keys, and student work.

After the semester withdrawal deadline passed, the Office of General Education Assessment requested the class rosters for all General Education courses in Mathematics from the Office of the Registrar and systematically selected every fourth student for assessment from the roster. Instructors of all General Education courses in Mathematics were sent a list of students selected for the assessment along with detailed instructions requesting that instructors provide a copy of one assignment along with the ungraded responses for the selected students to be sent via email to the Assessment Coordinator. In contrast from previous Mathematics assessments, faculty were also asked to provide answer keys based on feedback from past readers about the additional time it takes to calculate responses when determining “correctness” of student responses.

Student artifacts were collected and stored in an electronic repository and uploaded into the LiveText© assessment management system. A panel of faculty (tenured and tenure-track faculty, term faculty, and adjunct faculty), graduate teaching assistants, and REACH (Resources for Academic Achievement) Mathematics staff assessed student artifacts. The university’s Critical Thinking Rubric for Mathematics was applied to all student artifacts. Prior to the assessment reading, assessors were brought together for a four-hour training session coordinated by the Office of General Education Assessment. In the training session, the assessment process and context for General Education Assessment at the University of Louisville were presented. Faculty engaged in dissection and discussion of rubric criteria, and faculty assessors individually reviewed and scored benchmark sample assignments. Benchmarks were assignments selected to represent a wide range of content and skill development in order to give the assessors a baseline for measuring expectations of learning and evaluating student performance (Herman, Osmundson, & Dietel, 2010). Assessors then engaged in discussion about the benchmark assessment scores to share their rationales for why particular scores were selected. To highlight the reliability of the training scoring, the results from scoring benchmark samples for the Mathematics Rubric are provided in Table 1 and Figure 1.

Table 1

Results of Benchmark Sample 1 – Mathematics Rubric

Rubric Measures	<u>Clearly Evident (4)</u>	<u>Usually Evident (3)</u>	<u>Minimally Evident (2)</u>	<u>Not Evident (1)</u>
M1	73.7%	26.3%	0	0
M2	73.7%	26.3%	0	0
M3	47.4%	47.4%	5.3%	0
M4	52.6%	47.4%	0	0

Results of Benchmark Sample 2 – Mathematics Rubric

Rubric Measures	<u>Clearly Evident (4)</u>	<u>Usually Evident (3)</u>	<u>Minimally Evident (2)</u>	<u>Not Evident (1)</u>
M1	100.0%	0	0	0
M2	68.4%	31.6%	0	0
M3	31.6%	68.4%	0	0
M4	0	94.7%	5.3%	0

Results of Benchmark Sample 3 – Mathematics Rubric

Rubric Measures	<u>Clearly Evident (4)</u>	<u>Usually Evident (3)</u>	<u>Minimally Evident (2)</u>	<u>Not Evident (1)</u>
M1	42.1%	57.9%	0	0
M2	47.4%	52.6%	0	0
M3	0	77.8%	22.2%	0
M4	42.1%	52.6%	5.3%	0

Table 1 Continued

Results of Benchmark Sample 4 – Mathematics Rubric

Rubric Measures	<u>Clearly Evident (4)</u>	<u>Usually Evident (3)</u>	<u>Minimally Evident (2)</u>	<u>Not Evident (1)</u>
M1	100.0%	0	0	0
M2	94.7%	5.3%	0	0
M3	100.0%	0	0	0
M4	52.6%	47.4%	0	0

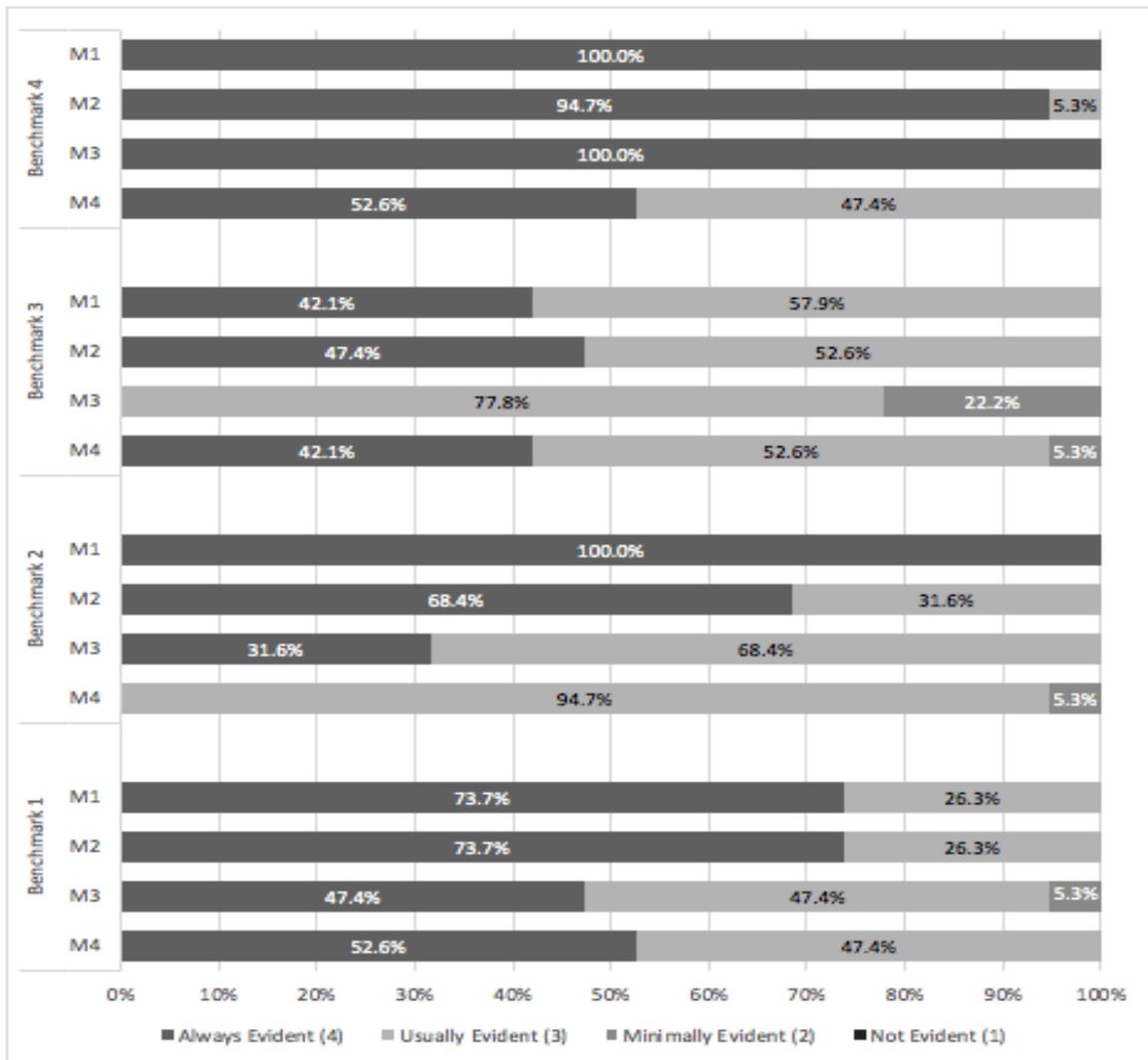


Figure 1. Results of Benchmark Sample Assessments for Mathematics

At the start of the assessment reading day, each faculty assessor was assigned a username and password for one of three LiveText© accounts and a list of courses and sections to assess. Three readers assessed each artifact so that scores could be compared across assessors for reliability purposes.

Data Collection Overview

The enrollment for Mathematics General Education courses in Spring 2018 was approximately 1354 students after the withdraw deadline. The Office of General Education Assessment requested artifacts from every fourth student and all students from select dual-credit courses. A total of 331 artifacts were received and determined to be eligible for assessment.

Summary of Assessment Data

For the assessment of Mathematics outcomes, 331 student artifacts were assessed by faculty and graduate teaching assistants from the College of Arts & Sciences and the J. B. Speed School of Engineering, as well as REACH (Resources for Academic Achievement) Mathematics staff, using the Mathematics Critical Thinking Rubric. A summary of results from the Mathematics assessment is provided in Table 2 and Figure 2.

The criterion for both the Mathematics Rubric was set by the General Education Assessment Coordinator and the General Education Curriculum Committee Assessment Subcommittee at 60% of artifacts to score at a 3 or 4, indicating that at least 60% demonstrate performance at either the “usually evident” or “clearly evident” level. The criterion was met for M1 (correctly interprets mathematical information), fell just short for M2 (applies mathematical models to solve problems) and M3 (represents mathematical information), and was not met for M4 (provides complete, reasonable, and correct answers).

Table 2

Summary Results for Mathematics Assessment

Number of Artifacts Scored at Each Rubric Level

	M1	M2	M3	M4
Clearly Evident (4)	357	321	317	195
Usually Evident (3)	266	246	251	256
Minimally Evident (2)	284	315	311	358
Not Evident (1)	86	111	114	174

Table 2 Continued

Percentage of Artifacts Scored at Each Rubric Level

	M1	M2	M3	M4
Clearly Evident (4)	36.0%	32.3%	31.9%	19.6%
Usually Evident (3)	26.8%	24.8%	25.3%	26.8%
Minimally Evident (2)	28.6%	31.7%	31.3%	36.1%
Not Evident (1)	8.7%	11.2%	11.5%	17.5%
% Scored at 4 & 3	62.7%	57.1%	57.2%	46.4%

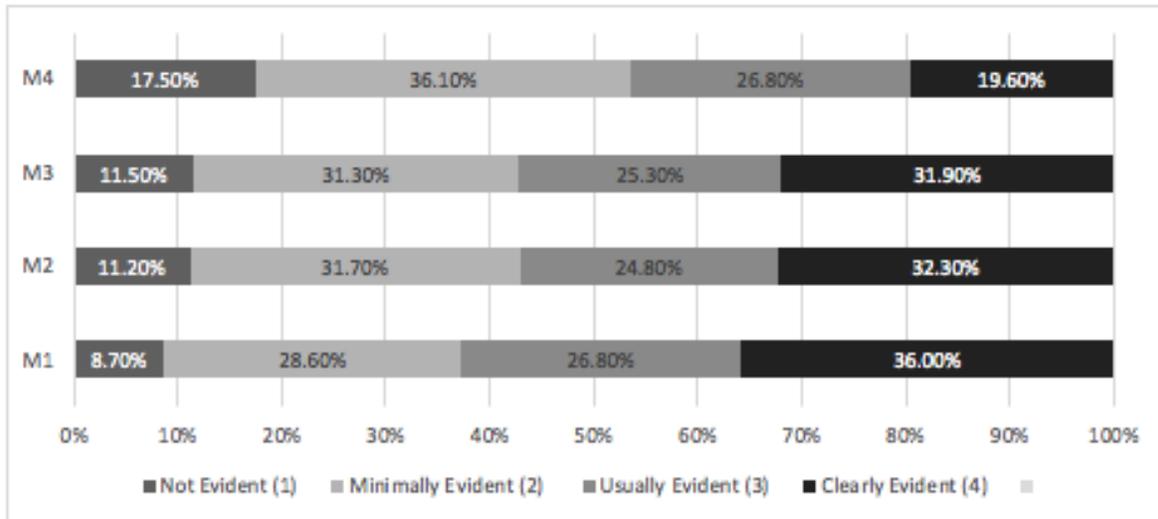


Figure 2. Summary Results for Mathematics Assessment

The mean, median, and mode for each rubric measure is provided in Table 3. The mode was at “Clearly Evident (4)” for the first three rubric measures and at “Minimally Evident (2)” for the fourth rubric measure.

Table 3

Mean, Median, and Mode for Mathematics Assessment

	M1	M2	M3	M4
Mean	2.90	2.78	2.78	2.49
Median	3	3	3	2
Mode	4	4	4	2

Inter-rater Reliability

Three separate readers assessed each student artifact. Table 4 displays the mean score for the three separate readings of all artifacts.

Table 4

Inter-rater Summary for Mathematics Assessment

	Assessor 1	Assessor 2	Assessor 3	Standard Deviation
M1	2.80	2.88	3.02	.11
M2	2.77	2.75	2.83	.04
M3	2.66	2.87	2.80	.11
M4	2.41	2.56	2.49	.08

In addition to the descriptive statistics, Table 5 provides multiple measures of inter-rater reliability. The percentage agreement value was calculated to determine the percentage of artifacts for which all three assessors scored at the either the same or within one performance level. Values for *Total Agreement* provided in Table 5 represent the percentage of artifacts for which all three assessors selected the same score (e.g., Assessors 1, 2, and 3 all selected 3). *Agreement (within 1 level)* represents the percentage of artifacts for which all three assessors scored the artifact at the same performance level or within one level (e.g., Assessor 1 selected a score of 3, Assessor 2 selected a score of 2, and Assessor 3 also selected a score of 2).

In addition to percentage agreement, a one-way, average-measures intra-class correlation coefficient (ICC) was calculated to assess inter-rater reliability. ICC coefficients between .75 and 1.00 are considered excellent, .60 to .74 considered good, .40 to .59 fair, and below .4 is considered poor (Cicchetti, 1994). Based upon these criteria, inter-rater reliability was excellent for all rubric measures.

Table 5

Inter-rater Reliability for Mathematics Assessment

Competency Measure	Total Agreement	Agreement (within 1 level)	ICC	95% Confidence Interval
M1	36.6%	87.0%	.86	(.83-.88)
M2	37.8%	85.5%	.87	(.84-.89)
M3	32.0%	85.5%	.85	(.82-.88)
M4	41.4%	88.2%	.88	(.85-.90)

Summary and Plan for Improvement

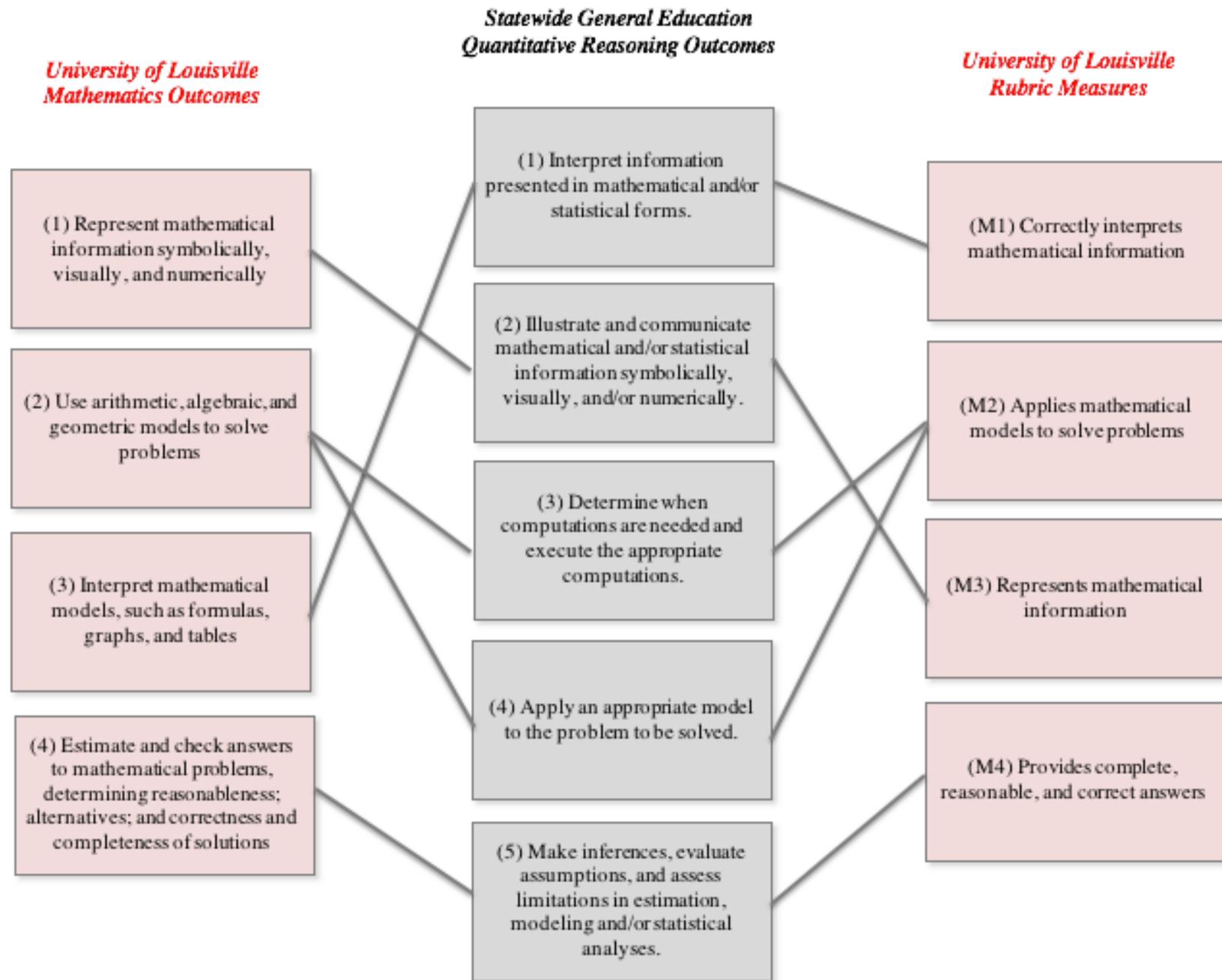
The Spring 2018 assessment of Mathematics concluded the final cycle under the university's current General Education Program. The university's new general education program, Cardinal Core, became effective in summer 2018. Under Cardinal Core, the Mathematics content area has been replaced by Quantitative Reasoning.

While the results for this assessment indicate that not all thresholds were met, the mode was at "Clearly Evident" for three of four rubric measures. Feedback from assessment readers indicated that the lower performance on the fourth rubric measure was due to the type of assignments that were collected for assessment. Assignments contained multiple problems, introducing more opportunity for student error.

References

- Association of American Colleges and Universities. (2009). *Oral communication VALUE Rubric*. Retrieved from <https://www.aacu.org/value/rubrics/oral-communication>
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment, 6*(4), 284-290. doi: 10.1037/1040-3590.6.4.284
- Fletcher, I., Mazzi, M., & Nuebling, M. (2011). When coders are reliable: The application of three measures to assess inter-rater reliability/agreement with doctor–patient communication data coded with the VR-CoDES. *Patient Education and Counseling, 82*(3), 341-345. doi:10.1016/j.pec.2011.01.004
- Hallgren, K. A. (2012). Computing inter-rater reliability for observational data: An overview and tutorial. *Tutorials in Quantitative Methods for Psychology, 8*(1), 23-34.
- Herman, J. L., Osmundson, E., & Dietel, R. (2010). *Benchmark Assessment for Improved Learning*. (AACC Report). Los Angeles, CA: University of California.

Appendix A: Outcomes and Assessment Measures Crosswalk



Appendix B

General Education Mathematics Syllabus Review (Spring 2018)

History of the Syllabus Review

In 2012, the General Education Syllabus Review Project was initiated to evaluate the congruence of general education course syllabi with the approved content-specific general education student learning outcomes. Specifically, it was designed to determine: (a) if the student learning outcomes stated in each course syllabus are congruent with the approved content-specific general education learning outcomes, and (b) if corresponding assessment methods are stated that support the approved content-specific general education learning outcomes.

In the spring of 2015, the GECC Assessment Subcommittee proposed that the Syllabus Review Project be incorporated into the existing General Education Assessment Project. Therefore, the syllabi from each content area will be collected and reviewed by the Office of General Education Assessment in alignment with the corresponding assessment cycle.

This report summarizes the review process and the results of the syllabi review for the Mathematics content area.

Review Process

The Provost requests that all faculty load their syllabi to Blackboard® each semester. These syllabi are then available through the university's course catalog system. For the purpose of this review, the Office of General Education Assessment collected all Mathematics syllabi that were loaded to Blackboard in Spring 2018.

The review of syllabi sought to answer two questions:

- 1) Does the syllabus contain the content specific general education learning outcomes approved for the course?
- 2) Are assessment methods stated that support the content-specific general education learning outcomes approved for the course?

An evaluation of the congruence between the listed assessment methods with the content specific approved general education learning outcomes was not conducted when a reviewer determined that the syllabus does not contain a statement of the approved content specific general education learning outcomes.

Mathematics

The syllabus review included syllabi from 56 of the 57 Mathematics General Education course sections offered in the Spring of 2018 resulting in a 98.2% sample. Appendix Table 1 provides a

breakdown of the number of General Education of syllabi available, the number of syllabi with the outcomes stated, and the number of syllabi that also included the assessment methods.

Table 1.

Mathematics Sample

	<u>Syllabi Available</u>	<u>Outcomes Listed in Syllabus</u>	<u>Assessment Method</u>
Math Courses	56 (98.2%)	37 (64.9%)	24 (64.9%)

The review of the 56 General Education Mathematics syllabi identified 37 syllabi (64.9%) containing the content-specific general education learning outcomes approved for the course. Further review of the 37 syllabi containing the General Education Outcomes revealed that 24 syllabi (64.9%) also listed the assessment methods for the General Education Outcomes.