**Action Plan: Responding to Fall 2015**

**General Education Assessment Results**

**October 11, 2016**

**Joshua Fausty**

**Professor of English**

**Director of General Education**

**Fall 2016 Action Plan Goal:** Increase the percentage of students meeting Tier-level University-Wide Student Learning Goals on end-of-semester signature assignments by 10% over Fall 2015.

**Overview**

|  |  |  |
| --- | --- | --- |
| **Critical Thinking, Tier 1** | | |
|  | Tier 1 or Above  (Fall 2015) | Goal for Fall 2016 |
| Average Across Dimensions | 58% | 68% |
|  |  |  |  |
| **Oral Communication, Tier 1** | | |
|  | Tier 1 or Above  (Fall 2015) | Tier 1 or Above  (Goal for Fall 2016) |
| Average Across Dimensions | 60% | 70% |
|  |  |  |
| **Written Communication, Tier 1** | | |
|  | Tier 1 or Above  (Fall 2015) | Tier 1 or Above  (Goal for Fall 2016) |
| Average Across Dimensions | 58% | 68% |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Critical Thinking, Tier 2** | | | | | |
|  | | Tier 2 or Above  (Fall 2015) | | Tier 2 or Above  (Goal for Fall 2016) | |
| Average Across Dimensions | | 29% | | 39% | |
|  | | | | | |
| **Oral Communication, Tier 2** | | | | | |
|  | | Tier 2 or Above  (Fall 2015) | | Tier 2 or Above  (Goal for Fall 2016) | |
| Average Across Dimensions | | NA | | 30% | |
|  |  | |  | |
| **Written Communication, Tier 2** | | | | | |
|  | | Tier 2 or Above  (Fall 2015) | | Goal Tier 2 or Above  (Goal for Fall 2016) | |
| Average Across Dimensions | | 19% | | 29% | |

In order to achieve the Fall 2016 Action Plan Goal, the following strategies will be implemented:

Strategy 1: Implement Robust Faculty Development Initiative (Face-to-Face Events and Electronic)

Strategy 2: Implement Professional Learning Communities (Ongoing Praxis Among Faculty)

Strategy 3: Encourage Transparent Assignment Design (Purpose [Knowledge/Skills], Tasks, Criteria)

Strategy 4: Strengthen Targeted Pedagogical Strategies (Scaffolding Skills-Building in the Disciplines)

Strategy 5: Improve Data Collection (Submission Rates)

Strategy 6: Improve Inter-Rater Reliability (of Assessment Team)

Strategy 7: Pilot Assessment by Faculty (Faculty Complete Rubrics in Parallel with Assessment Team)

Strategy 8: Universalize Closing-the-Loop Efforts (Faculty Respond to Assessment Results)

*See below for details.*

**Strategy Details**

Strategy 1: Implement Robust Faculty Development Initiative (Face-to-Face Events and

Electronic)

* + “Teaching in Gen Ed,” a two-part faculty development event for Gen Ed instructors, will bring colleagues together to share strategies for supporting student success and provide answers to questions about the new program. Part 1 will address early-to-mid-semester concerns; Part 2 will focus on late-semester procedures and strategies for finishing up strong. The following topics will be covered:

**Part 1**

• The “Modes of Inquiry” approach to student learning across the disciplines

• Supporting student achievement of the University-wide student learning goals (UWSLGs) withoutsacrificing discipline-specific course objectives

• End-of-semester signature assignments: their role in the classroom and for programmatic assessment

• Transparent assignment design: a proven strategy for student success

• Scaffolding: building up to end-of-semester signature assignments

**Part 2**

• The UWSLG rubrics: assessment instruments and teaching tools

• Navigating Tk20: ensuring student submissions and completing the rubrics

• Gen Ed assessment: “Closing the loop” to improve student success.

Strategy 2: Convene Professional Learning Communities (PLCs) (Ongoing Praxis Among

Faculty)

* + Instructors will be assigned to teaching-and-learning communities, non-evaluative professional learning communities according to Mode(s) of Inquiry and UWSLGs covered in their courses.
  + Break-out sessions will be held at the Teaching in Gen Ed workshops where the PLCs will begin their collaborative work to address the Fall 2015 Gen Ed assessment results.
  + Instructors will join an online portal to facilitate communication, collaboration, and continued faculty development between in-person workshops (and to accommodate those unable to attend in-person workshops).
  + PLCs will exchange assignment plans and signature assignment drafts for discussion and feedback via the online portal.
  + PLCs will meet, virtually and in person, to discuss planning, progress, and strategies for improving the percentage of students who meet tier-level UWSLGs.
  + The PLCs will convene again at Part 2 of “Teaching in Gen Ed” to discuss and prepare for end-of-semester SA-related activities including drafting, revision, completion, submission, and scoring.

Strategy 3: Encourage Transparent Assignment Design (Purpose [Knowledge/Skills], Tasks,

Criteria)

* + Provide faculty development on principles of transparent assignment design to support student success at demonstrating UWSLGs.
  + Encourage faculty to adopt transparent design in their courses, especially on signature assignments—that is, to include clear and accessible information on the purposes of an assignment (knowledge and skills), the required tasks, and the criteria for success (including outcomes associated with the two UWSLGs covered in the course).
  + Emphasize importance of including all learning outcomes for covered UWSLGs in SA prompts (transparent assignment design should help reduce the number of NAs, which signify that a given outcome was not required by an assignment prompt.
  + Provide faculty development on curricular planning so that SAs are the culmination of a developmental progression of skills building throughout the course.

Strategy 4: Emphasize Development of Targeted Pedagogical Strategies (Scaffolding Skills-

Building in the Disciplines)

* + Share Fall 2015 assessment results with faculty and PLCs.
  + Emphasize need to include all UWSLGs covered in a given course on SAs.
  + Emphasize that UWSLGs are comprised of specific student learning outcomes to be covered during the semester and demonstrated on end-of-semester SA.
  + Provide professional development on scaffolding of skills-development and the need to build skills gradually by assigning increasingly sophisticated work on a continuum that spans the semester leading up to the SA.
  + Develop and discuss scaffolding and other strategies for improving student success on all outcomes, especially those of greatest concern.

Strategy 5: Improve Data Collection (Submission Rates)

* + Require students to submit SAs exclusively via Tk20.
  + Hold mid- and late-semester training workshops for students and faculty on navigating Tk20 and submitting SAs (begun in Fall 2016; continued in Spring 2016; to be continued in Fall 2016 with expansion to include NJCU Hub sessions and online video training).
  + Hold late-semester drop-in sessions for students and faculty on submitting SAs to Tk20 (begun in Fall 2016; continued in Spring 2016; to be continued in Fall 2016 with expansion to include NJCU Hub sessions and online video training).

Strategy 6: Improve Inter-Rater Reliability (of Assessment Team)

* + Expand assessment team training work to include double the number of sample SAs from actual student submissions as completed for Fall 2015.
  + Hold norming sessions for subgroups of the assessment team who work on each of the UWSLGs.
  + Provide average scores from among all assessment team members.
  + Provide opportunity for assessment team members to compare scores on actual student submissions.
  + Hold discussions to establish agreement on scoring standards and rationales including especially the proper use of the NA designation.

Strategy 7: Assessment of SAs by Faculty Instructors (Pilot Faculty Completing Rubrics in

Parallel to Ongoing Assessment Team Activities)

* + Faculty instructors will engage in the assessment of their own students’ learning in Gen Ed in a Fall 2016 pilot process to be implemented in parallel to ongoing assessment team work.
  + Faculty will be assured that the purpose of direct assessment of student learning is to inform faculty development initiatives, improve curriculum and instruction, and enhance student success—not to assess faculty teaching performance. It is essential that they understand the proper use of assessment results and know their honest participation in the assessment of their own students’ learning will not be used against them.
  + Faculty will be trained in the use of the program assessment rubrics to score their own students’ coursework.
  + Faculty will be asked to complete assessment rubrics on Tk20 for their own students’ signature assignments.
  + Data resulting from the work of the assessment team will be kept separate from data resulting from the assessment work of Gen Ed instructors completing rubrics for their own students’ SAs.
  + Data collected in the pilot will be reportable only in the aggregate.

Strategy 8: Universalize Closing-the-Loop Efforts (Faculty Respond to Assessment Results)

* + Direct knowledge of assessment results for their own classes will empower faculty to design and implement regular targeted adjustments to curriculum and instruction to improve student outcomes.
  + Increased faculty involvement in assessment will improve their sense of connection to the assessment process and have a greater investment in the results.
  + Faculty involvement in assessment activities will support effective closing-the-loop activities, given that such activities—designed in response to assessment results—must be undertaken by faculty, in the classroom, in response to assessment results.

**Fall 2015 General Education Assessment Results**

The above action plan responds to Fall 2015 assessment results for Critical Thinking and Problem-Solving, Oral Communication, and Written Communication, the main findings of which are presented in brief below. A full report on the Fall 2015 assessment results is available on GothicNet.

**Critical Thinking and Problem-Solving Results**

Of Note

* In Tier 1 courses, students did best on *Topic Identification*, with 80% at or above Tier 1. Although *a priori* targets for performance on dimensions were not set, 80% is a common target for program-level student learning outcomes assessment.
* Also in Tier 1, the dimension related to *Influence of Context* appeared to be most challenging, with only 40% at or above Tier 1. It also has the highest percentage below benchmark in Tier 1 and Tier 2.
* In Tier 2 courses, *Topic Identification* again showed the strongest performance, with 52% at or above Tier 2. The number of artifacts reviewed at this tier was limited.
* Juror agreement is poor, ranging from 28% to 44% across dimensions. In some cases, there were 3- or 4-point differences in ratings.
* Between 12% and 16% of jurors selected NA for each of the dimensions.

Critical Thinking Artifact Ratings, Tier 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dimension | Tier 3 | Tier 2 to < Tier 3 | Tier 1 to < Tier 2 | Benchmark to < Tier 1 | < Benchmark | Students |
| Topic identification, management | 0% | 27% | 54% | 15% | 5% | 164 |
| Explanation of issues | 0% | 16% | 54% | 25% | 5% | 166 |
| Evidence selecting, using information to investigate a point of view/conclusion | 1% | 8% | 43% | 40% | 8% | 166 |
| Influence of context, assumptions | 1% | 6% | 34% | 47% | 13% | 163 |
| Student's position (perspective, thesis/hypothesis) | 0% | 8% | 41% | 44% | 7% | 166 |
| Theoretical framework, approach | 1% | 15% | 47% | 29% | 8% | 165 |
| Conclusions and related outcomes | 1% | 6% | 47% | 38% | 9% | 165 |

Scores are averaged across jurors

|  |  |  |
| --- | --- | --- |
| Dimension | Tier 1 or Above | Below Tier 1 |
| Topic identification, management | 80% | 20% |
| Explanation of issues | 70% | 30% |
| Evidence selecting, using information to investigate a point of view/conclusion | 52% | 48% |
| Influence of context, assumptions | 40% | 60% |
| Student's position (perspective, thesis/hypothesis) | 49% | 51% |
| Theoretical framework, approach | 62% | 38% |
| Conclusions and related outcomes | 53% | 47% |

Critical Thinking Artifact Ratings, Tier 2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dimension | Tier 3 | Tier 2 to < Tier 3 | Tier 1 to < Tier 2 | Benchmark to < Tier 1 | < Benchmark | Students |
| Topic identification, management | 0% | 52% | 38% | 10% | 0% | 21 |
| Explanation of issues | 0% | 36% | 9% | 55% | 0% | 22 |
| Evidence selecting, using information to investigate a point of view/conclusion | 0% | 24% | 43% | 33% | 0% | 21 |
| Influence of context, assumptions | 0% | 16% | 37% | 32% | 16% | 19 |
| Student's position (perspective, thesis/hypothesis) | 0% | 14% | 36% | 50% | 0% | 22 |
| Theoretical framework, approach | 0% | 41% | 36% | 18% | 5% | 22 |
| Conclusions and related outcomes | 0% | 17% | 38% | 33% | 13% | 24 |

Scores are averaged across jurors

|  |  |  |
| --- | --- | --- |
| Dimension | Tier 2 or Above | Below Tier 2 |
| Topic identification, management | 52% | 48% |
| Explanation of issues | 36% | 64% |
| Evidence selecting, using information to investigate a point of view/conclusion | 24% | 76% |
| Influence of context, assumptions | 16% | 84% |
| Student's position (perspective, thesis/hypothesis) | 14% | 86% |
| Theoretical framework, approach | 41% | 59% |
| Conclusions and related outcomes | 17% | 83% |

Percentage NA (Critical Thinking, Tier 1 and 2)

|  |  |
| --- | --- |
| Dimension | 348 Assessments |
| Topic identification, management | 15% |
| Explanation of issues | 12% |
| Evidence selecting, using information to investigate a point of view/conclusion | 14% |
| Influence of context, assumptions | 16% |
| Student's position (perspective, thesis/hypothesis) | 13% |
| Theoretical framework, approach | 15% |
| Conclusions and related outcomes | 13% |

**Oral Communication Results**

Of Note

* Only Tier 1 courses were included in these analyses; there were insufficient Tier 2 courses for inclusion.
* Over three quarters of oral presentations were scored at or above Tier 1 on the *Language* dimension, the highest of the five dimensions. This is close to the common target of 80%.
* An average of 20% of oral presentations fell below the benchmark level; this is significantly lower than the percentage below benchmark on the other learning goals.
* The *Supporting Material* dimension had the most juror NAs. In addition, for this dimension, there was a 3- or 4-point difference in scoring for 26% of juror pairs, which is quite high. Strangely, the percent of jurors with perfect agreement is quite good, but the percent within one point is not so good. There may have been confusion among jurors about what was meant by supporting material in the context of an oral presentation. This dimension requires further investigation.
* Oral Communication has a higher percentage below benchmark than the other learning goals, with *Delivery* the highest at 29% of students below benchmark.

Oral Communication Artifact Ratings, Tier 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dimension | Tier 3 | Tier 2 to < Tier 3 | Tier 1 to < Tier 2 | Benchmark to < Tier 1 | < Benchmark | Students |
| Organization | 2% | 12% | 50% | 20% | 16% | 123 |
| Language | 2% | 15% | 59% | 15% | 10% | 123 |
| Delivery | 1% | 11% | 41% | 17% | 29% | 123 |
| Supporting material | 2% | 14% | 33% | 27% | 24% | 120 |
| Central message | 2% | 13% | 44% | 20% | 21% | 123 |

Scores are averaged across jurors

|  |  |  |
| --- | --- | --- |
| Dimension | Tier 1 or Above | Below Tier 1 |
| Organization | 64% | 36% |
| Language | 76% | 24% |
| Delivery | 54% | 46% |
| Supporting material | 49% | 51% |
| Central message | 59% | 41% |

Percentage NA (Oral Communication, Tier 1)

|  |  |
| --- | --- |
| Dimension | 253 Assessments |
| Organization | 15% |
| Language | 14% |
| Delivery | 16% |
| Supporting material | 21% |
| Central message | 16% |

**Written Communication Results**

Of Note

* Students in both Tier 1 and Tier 2 have highest scores in *Mechanics*, with 64% in Tier 1 and 32% in Tier 2 at or above the appropriate tier.
* The sample size for Tier 2 is small.
* The *Sources* dimension may require further examination – 12% of artifacts in Tier 1 and 20% in Tier 2 were below benchmark. In addition, there were relatively few NAs, compared to the two other learning goals, except for *Sources* where 20% of jurors indicated NA.
* As with the other two learning goals, juror agreement was low, ranging from 30% for *Context* to 43% for *Mechanics.*

Written Communication Artifact Ratings, Tier 1

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dimension | Tier 3 | Tier 2 to < Tier 3 | Tier 1 to < Tier 2 | Benchmark to < Tier 1 | < Benchmark | Students |
| Context/purpose | 0% | 15% | 47% | 32% | 5% | 169 |
| Content development | 0% | 11% | 47% | 39% | 4% | 171 |
| Conventions | 0% | 7% | 45% | 40% | 7% | 168 |
| Sources | 0% | 6% | 48% | 34% | 12% | 161 |
| Mechanics | 1% | 7% | 56% | 33% | 4% | 171 |

Scores are averaged across jurors

|  |  |  |
| --- | --- | --- |
| Dimension | Tier 1 or Above | Below Tier 1 |
| Context/purpose | 63% | 37% |
| Content development | 58% | 42% |
| Conventions | 52% | 48% |
| Sources | 54% | 46% |
| Mechanics | 64% | 36% |

Written Communication Artifact Ratings, Tier 2

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Dimension | Tier 3 | Tier 2 to < Tier 3 | Tier 1 to < Tier 2 | Benchmark to < Tier 1 | < Benchmark | Students |
| Context/purpose | 0% | 23% | 50% | 27% | 0% | 30 |
| Content development | 0% | 16% | 58% | 26% | 0% | 31 |
| Conventions | 0% | 16% | 45% | 39% | 0% | 31 |
| Sources | 0% | 8% | 48% | 24% | 20% | 25 |
| Mechanics | 0% | 32% | 55% | 10% | 3% | 31 |

Scores are averaged across jurors

|  |  |  |
| --- | --- | --- |
| Dimension | Tier 2 or Above | Below Tier 2 |
| Context/purpose | 23% | 77% |
| Content development | 16% | 84% |
| Conventions | 16% | 84% |
| Sources | 8% | 92% |
| Mechanics | 32% | 68% |

Percentage NA (Written Communication, Tier 1 and 2)

|  |  |
| --- | --- |
| Dimension | 383 Assessments |
| Context/purpose | 10% |
| Content development | 5% |
| Conventions | 7% |
| Sources | 20% |
| Mechanics | 5% |

**Summary Data**

The following three tables show how many students achieved the Tier 1 level, fell below the Tier 1 level, and fell below Benchmark for the three learning goals assessed in Fall 2015. Note that the percentage “Below Benchmark” is also included in the percentage for “Below Tier 1” (which explains why the total of each row adds up to more than 100%):

|  |  |  |  |
| --- | --- | --- | --- |
| **Critical Thinking, Tier 1** | | | |
|  | Tier 1 or Above | Below Tier 1  (Also counted in “Below Benchmark”) | Below Benchmark  (Also counted in “Below Tier 1”) |
| Average Across Dimensions | 58% | 42% | 8% |
|  |  |  |  |
| **Oral Communication, Tier 1** | | | |
|  | Tier 1 or Above | Below Tier 1  (Also counted in “Below Benchmark”) | Below Benchmark  (Also counted in “Below Tier 1”) |
| Average Across Dimensions | 60% | 40% | 20% |
|  |  |  |  |
| **Written Communication, Tier 1** | | |  |
|  | Tier 1 or Above | Below Tier 1  (Also counted in “Below Benchmark”) | Below Benchmark  (Also counted in “Below Tier 1”) |
| Average Across Dimensions | 58% | 42% | 6% |

Students in Tier 1 classes achieved the Tier 1 level 58-60% of the time in all three learning goals. Since Fall 2015 was our first semester assessing these learning goals, there is no comparative data; however, a common target for this type of assessment is 80%. In Fall 2015, students reached the Tier 1 level 80% of the time in only one dimension, “Topic Identification and Management” under Critical Thinking. The next highest level of achievement was in “Language” under Oral Communication, for which 76% of students achieved the Tier 1 level for the course. Although achievement was comparable across the three learning goals in Tier 1, a large percentage (20%) of students fell below the benchmark in Oral Communication. This is considerably more than the 6-8% under benchmark for the other two learning goals. All dimensions under Oral Communication had high percentages below benchmark (10-29%). “Delivery” had the highest percentage under benchmark for Oral Communication. “Supporting Material” showed the smallest percentage of students reaching the Tier 1 level for Oral Communication; it also showed more NAs than other dimensions.

In contrast, 76% of students in Tier 1 achieved the Tier 1 level for “Language” under Oral Communication. “Language” also had 80% agreement within one point among jurors, the highest of the dimensions of Oral Communication. The combined Oral Communication results indicate that “Language” skills are not the main issue leading to the high percentage of students falling below Benchmark for Oral Communication; rather, the “Supporting Material” dimension stands out as a challenge area deserving of greater attention.

When all three learning goals are averaged together, 62% of students in Tier 1 and Tier 2 classes achieved the Tier 1 level. Students in Tier 2 classes were more likely to meet Tier 2 standards than students in Tier 1. However, as the tables below reveal, students in Tier 2 courses did not meet Tier 2 standards at a high rate.

|  |  |  |  |
| --- | --- | --- | --- |
| **Critical Thinking, Tier 2** | | | |
|  | Tier 2 or Above | Below Tier 2  (Also counted in “Below Benchmark”) | Below Benchmark  (Also counted in “Below Tier 2”) |
| Average Across Dimensions | 29% | 71% | 5% |
|  |  |  |  |
| **Written Communication, Tier 2** | | | |
|  | Tier 2 or Above | Below Tier 2  (Also counted in “Below Benchmark”) | Below Benchmark  (Also counted in “Below Tier 2”) |
| Average Across Dimensions | 19% | 81% | 5% |

The two tables above show that 19%-29% of students in Tier 2 courses achieved the Tier 2 outcomes in Fall 2015, the first generation of the new program and a semester in which all students were first-semester freshmen with no previous college-level experience. (Transfer students were not enrolled in General Education courses in Fall 2015.)

While the percentage of assessments assigned NA was generally comparable among most dimensions of the three learning goals (5-15%), two dimensions stand out. For Oral Communication, the “Supporting Material” dimension was assigned NA on 20% of assessments; while for Written Communication, the “Sources” dimension was assigned NA on 21% of assessments. It is noteworthy that these two dimensions measure similar skills—the use of external sources (or material) to support a position, perspective, or argument—through different delivery methods. The high percentage of NAs assigned on these dimensions suggests that 1) instructors may need to convey clearer expectations to students about their use of sources and supporting material; and 2) assessment team members may require feedback about what counts as an effective use of “supporting material” on an Oral Communication assignment, and as an effective use of “sources” on a Written Communication assignment. These two dimensions represent an area of particular concern for faculty development work on more effective assignment design *and* more effective assessment/scoring (i.e. juror training).

Among assignments scored for “Sources,” results were lower than on other dimensions for Tier 2 Written Communication (only 8% achieved Tier 2-level). Among assignments scored for “Supporting Material,” results were lower than on other dimensions for Tier 1 Oral Communication (49% achieved Tier 1-level). “Supporting Material” was generally weakest as the only dimension on which more than half of the artifacts in Tier 1 courses rated below the Tier 1-level. This suggests that, when assignment questions/prompts require students to use sources and supporting evidence, students have difficulty doing so. In addition to raising concerns about assignment design on these two dimensions, the data point to a need to find more effective ways to advance student achievement of these outcomes. As stated above, “Sources” and “Supporting Materials” represent priority areas for targeted faculty development.

After “Supporting Materials” and “Sources,” the dimensions with the next highest percentage of NAs were “Influence of context, assumptions” for Critical Thinking and “Delivery” and “Central message” for Oral Communication. These dimensions also represent areas of focus for improved assignment design and juror training; they each received NAs on 16% of scored assessments.

The high number of NAs assigned across dimensions highlights the need for instructors to clarify, on assignment questions/prompts, the specific requirements for each dimension of the learning goals to ensure that students are fully aware of requirements and expectations and are thus better prepared to demonstrate the appropriate outcomes. Greater specificity on assignments, coupled with targeted coursework designed to improve the requisite skills, should help maximize student achievement of the outcomes while minimizing the number of NAs assigned. In addition to transparent assignment design, more effective communication of, and training in, these specific learning goals and expectations throughout the semester will support program-wide student achievement of the University-wide student learning goals.

For the first semester of implementation, the following inter-rater reliability results are somewhat hopeful: juror pairs reported scores *within one point of each other* about 70-90% of the time for every dimension of all three learning goals except “Supporting Material” under Tier 1 Oral Communication. Furthermore, average juror agreement *within one point* across Tiers 1 and 2 was 82% for Critical Thinking; 84% for Written Communication; and 69% for Oral Communication—and the latter goes up to 73% if “Supporting Materials” is excluded (“Supporting Materials” had by far the lowest juror agreement at 54%). These inter-rater reliability results provide a clear starting point from which to improve. That said, a key finding of the Fall 2015 assessment report is that work remains to be done to increase juror agreement. A few areas of significant disagreement need to be addressed including the number of 3- and 4-point disagreements and the tendency of jurors to disagree among the various dimensions of the learning goals. Perfect juror agreement was achieved on average across the seven dimensions of Critical Thinking (Tiers 1 and 2) 28%-37% of the time; for Oral Communication (Tier 1 only), the range for perfect agreement was 35%-58%; and for Written Communication, the range for perfect agreement was 30%-43%. While across-the-board achievement of perfect agreement is unrealistic, it is not unreasonable to expect that continued assessment practice, as well as reflection and communication among jurors, instructors, and students, will have a positive impact on future results. In addition to the primary work of improving student learning outcomes, efforts to increase inter-rater reliability must be a priority.

**Strategy 7: the Fall 2016 Pilot (Rationale)**

Assessment team participation has been identified as a valuable experience for instructors who have participated—one akin to participation in an intensive faculty development program. Working with colleagues from across the four colleges and dozens of departments, and reviewing student work done in many different courses, takes instructors outside their classrooms and comfort zones, contributes to their knowledge and experience, and allows them to contribute to the enhancement of the knowledge of their colleagues. This can have a positive impact on everyone’s success by helping to ensure that expectations for student achievement are both reasonable and, increasingly, shared. It helps to build a wider culture of learning. Instructor participation in assessment activities also mitigates the need for greater communication and understanding among two groups, instructors and assessment team members, by effectively making them a single group with shared interests and experiences. A greater number of instructors scoring signature assignments might even lead to increased inter-rater reliability.

Unfortunately, it has been a challenge to bring a greater number of instructors onto the assessment team. Furthermore, doing so may not be the most efficient way to increase instructor participation in assessment activities. A more direct and useful strategy would be to ask all instructors to participate in the assessment of their own students’ work, and to focus programmatic attention on providing opportunities for all faculty involved in teaching and assessment to come together to discuss results and strategize about ways to improve outcomes. While the first phase of assessment activities, relying on the assessment team, has provided a baseline of data from which to move forward, it is also true that few instructors have been involved, and the extent to which instructors will be likely to participate in closing-the-loop activities based on data they had little to no role in producing remains to be seen. In addition, the lack information on which students did well and which poorly on each of the outcomes and learning goals renders the assessment results less meaningful to faculty who may be likely to assume that the aggregated results do not accurately reflect the performance of their own students. A method of data collection that allows for disaggregating results so that faculty may see how their own students perform, and how they fit into the larger programmatic picture, will enhance the value of assessment activities and lead to more relevant and actionable results.

Having instructors score their own students’ work using the program rubrics will dramatically reduce the complexity involved in having an outside assessment team, the great majority of whom are not teaching in the program, score the work of students. In addition, transferring the main scoring practice to the instructors will make implementation of closing-the-loop activities more straightforward as the instructors will have more intimate knowledge of the results and their true significance due to their direct involvement in the process scoring the work of their own students.

During programmatic assessment in Fall 2015 and Spring 2016, the names and identities of students and instructors were detached from the assessment results of the students whose work was being assessed by the assessment team. Results for each learning goal were broken down according to the Tier level of the course in which the student did the assignment, but no further information could be gleaned from the results. This meant that any results were aggregate and could not be disaggregated beyond the Tiers. For example, we have data on the number of students in Tier 1 and Tier 2 classes who reached the Tier 1 or Tier 2 level of proficiency in each of the learning goals, and in each of the outcomes for each of the learning goals, but we are unable to access more detailed data, for example, to provide each instructor with a report on his or her own students’ results and a comparison to the relevant aggregate programmatic data. As a result, it has been impossible for instructors to design more targeted curricular or pedagogical interventions based on course-specific results in the context of program-wide results. The first phase of programmatic assessment thus severely limited the power of assessment to inform specific interventions to improve student learning. We will be able to learn a lot more from our assessment activities through the implementation of a few adjustments to the plan starting with a pilot that asks faculty to score their own students’ work using the program rubrics on Tk20.

In Fall 2016, instructors teaching General Education courses will be asked to complete assessment rubrics for end-of-semester signature assignments in their courses to assess their own students’ achievement of the outcomes associated with each of the two learning goals covered. Piloting individual instructor involvement in programmatic assessment will allow disaggregated data to be marshaled in service of curricular and pedagogical improvements to increase student success.

The Assistant Vice President for Institutional Effectiveness has agreed to implement a coding system to prevent raw data on individual student performance from being linked to individual faculty instructors in case of an accidental release of assessment data. The disaggregated data will be released on a very limited basis. In fact, only the Director of General Education and the Assistant Vice President for Institutional Effectiveness will have access to disaggregated data; and reports of disaggregated data on individual classes will be made exclusively to the faculty who taught those classes.

Director of General Education and the Assistant Vice President for Institutional Effectiveness have agreed to the following:

Statement on Disclosure of

Direct Assessment Data in General Education

Any and all direct assessment data collected in connection with the 2016-17 pilot program of instructor scoring of General Education Signature Assignments will not be made available to individuals who are in a position to make or influence personnel decisions about individual faculty participants. Furthermore, only aggregate data, with no cells small enough to identify individual faculty participants, will be shared publicly, with the exception of disaggregated data for specific course sections being shared exclusively with instructors of those specific course sections.

As the specified above, individual reports will be shared with faculty on their own students’ results, which will make it possible for those results to guide faculty responses to their students’ performance from semester to semester. This level of feedback is not possible in the first phase of the assessment plan, during which no data is currently available on individual student or course section performance.

In addition, the Director of General Education has consulted with the Provost and Dean of Arts and Sciences, both of whom have agreed that **data resulting from the direct assessment of student learning in General Education will not be used to evaluate faculty teaching performance.**The purpose of assessment is to inform faculty development initiatives, improve curriculum and instruction, and enhance student success in addition to providing data on programmatic success.

Instructors should feel comfortable getting involved in a process meant to improve teaching and learning and support their students’ success. The point is to find out how our students are doing and to help them do better, not to evaluate instructors. Instructors need to trust that their voluntary and honest evaluation of student performance will not be used against them—or they won’t participate honestly.

Faculty participation in the assessment of student learning should also be understood as voluntary to ensure its usefulness. **Gen Ed faculty will be *asked and encouraged* to complete rubrics on student achievement, *not required* to do so. To further emphasize that participation is voluntary, their participation, or lack thereof, will not be reported except in the aggregate.**

In Fall 2016, faculty scoring of signature assignments will parallel the assessment team’s continuing work. This will provide an opportunity to assess the assessment processes to determine what arrangement will best meet the needs of the program and support student achievement going forward.