***Q Discussions Overview***

Spring 2014 – Spring 2015

|  |
| --- |
| **General Education Goal Assessed** |
| This committee is charged with assessing the Q / QM / QS skills. From the Faculty Handbook:  *There are two parts to the Quantitative Skills component: a statistical part (QS), and a mathematical part (QM). Courses that satisfy the statistics requirement will carry a QS designation and should contain elementary statistics topics such as averages, standard deviation and other measures of dispersion, as well as interpretation of data, tables, graphs, and some probability. Courses that satisfy the mathematical requirement will carry a QM designation and must use a combination of algebraic, graphical, and numerical reasoning. Such courses should teach students how to translate problems into mathematical language, how to solve the mathematical problems, and how to interpret the solutions. Courses that carry a Q designation must fulfill the requirements for both the statistical (QS) and mathematical (QM) components.*  *Courses with quantitative skills components necessarily involve the use of appropriate technology.*  *Students have two options for fulfilling the Quantitative Skills component. They may either 1) complete a single course that carries the Q designation or 2) complete a course that fulfills a QS designation and complete a course that carries a QM designation.*  The challenge the committee faced was in converting the quantitative “inputs” to students presented in this definition to measurable student learning outcomes. In addition, finding a single measureable outcome for all the QM courses, MA 130, MA 100, and EB 210 was challenging, so the committee “split” the assessments. |
| **Development of the Q Assessment Plan, Spring 2014 Semester** |
| In late Fall 2013 and early Spring 2014, the call to faculty was made for participants on the committee to assess the Quantitative skills requirement of the General Education curriculum. Quantitative courses hold the Q, QM, and QS designations. As suggested by Juniata’s existing General Education Assessment plan, articulated in Spring 2013, Assistant Provost Gerald Kruse and Director of Institutional Research Carlee Ranalli were designated as co-chairs. Eventually, Kim Roth from the Mathematics Department, Matt Beaky from the Physics Department, Russ Shelley from the Music Department, and John Mumford, the Head Librarian, agreed to serve. Kim teaches courses with Q, QM, and QS designations and Matt teaches courses with the QM and QS designations. John was interested in serving in light of the information literacy component of the Q designation. Russ was a member of the ***Q Subcommittee*** of the Curriculum Faculty standing committee, which until Spring 2013 was charged with considering new courses proposed with some Quantitative designation. Jerry was the chair of this Q Subcommittee, until it was disbanded by the Curriculum Committee at the end of the Spring 2013 semester. The third member of the Q Subcommittee, Ron McLaughlin from the Psychology Department, declined the invitation to serve on this assessment committee, citing concerns over instructors of Quantitative courses undertaking the assessment of these same courses. It should be noted that Jerry’s faculty appointment is in the Mathematics and IT/CS departments, and he teaches courses with Q and QM designations.  The committee met twice a month during the Spring 2014 semester to design an assessment plan, which was to be implemented in the Fall 2014 semester.  Initially, the committee focused on  (1) understanding which courses are offered with some type of Quantitative designation  (2) when students satisfy their Quantitative requirement  (3) considering what student learning outcomes are associated with these classes.  The Juniata Associate for General Education Assessment during the 2013-2014 academic year, Tori Rehr, was unable to attend committee meetings due to a class conflict, but she was able to contribute to the work of the committee by doing research and analysis of the syllabi (<http://services.juniata.edu/faculty/syllabi/index.html?a=upload>) of Quantitative courses. A summary of this analysis follows:  Selected Q Course Objectives  • Analyzing and collecting data  • Critical thinking to statistically defend decisions  • Interdisciplinary nature of data  • Emphasis on higher-level skills  • Interpret and understand graphs and data  • Application with computer science and computer software  Selected QM Course Objectives  • Note: Course objectives were much more detailed with business and physics departments, with math course  objectives tending to focus more on topics to be covered rather than skill acquisition.  • Write to clearly identify problems and recommend solutions.  • Proficiency with computer spreadsheets  • Use tools of financial analysis  • Work through problems independently  • Enrichment of understanding of scientific outcome  • Apply mathematic theory to “real word problems”  Selected QS Course Objectives  • Ability to design experiments or surveys  • Apply and differentiate statistical models and tests  • Interpret graphs  • Visualization techniques for data  • Create “best estimate” using parameters and statistical techniques  • Interpreting data for public consumption and relevant data  • Promote development of problem solving ability  This analysis helped frame the committee’s discussion of student learning outcomes in Quantitative courses, leading to the specific assessments detailed below.  In an attempt to understand approximately when and how students in a graduating class might satisfy their Quantitative skills requirement, the eight semesters for a “typical” member of the class of 2013 were considered, Fall 2009 through Spring 2013. Each of those semesters is listed in the table below, along with the number of Quantitative courses taken by the 2013 graduates. To clarify reading the first row of the table, it does not necessarily mean that exactly 288 students took a Q course that Fall 2009 semester, but credits for 288 courses were earned during the Fall 2009 semester by the 2013 graduates.   |  |  | | --- | --- | | Semester | Number of Q courses taken by students who graduated in May 2013  (NOT number of students taking Q courses) | | 09/FA | 288 | | 10/SP | 328 | | 10/FA | 219 | | 11/SP | 160 | | 11/FA | 125 | | 12/SP | 111 | | 12/FA | 86 | | 13/SP | 77 |   The rest of the Quantitative course data generated by the Registrar’s Office and processed by this committee is available on the GenEd shared drive, in the folder \\jcshare\gea\Q-QM-QS .  In light of the number of Q courses taken in the early semesters by the “typical” student, the committee elected to focus on introductory level courses in this round of assessment, and so considered identifying instruments for these. In addition, these introductory courses are typically pre-requisites for the upper-level courses. The committee felt that the assessment of many upper-level courses would also overlap in departmental assessments. The committee decided to apply for approval from Juniata’s Institutional Research Board (IRB), where the ***Informed Consent*** form was used to determine the age of students participating. All students were expected to participate, but only the data from students 18 years or older was used.  To assess introductory QS courses, the committee considered an assessment from the Statistics Education Department at the University of Minnesota, the *Comprehensive Assessment of Outcomes in the first Statistics* *course* (CAOS). This assessment had many features which were attractive to the committee: it is free, nationally normed, 30 items administered online, 30-45 minutes to complete, items selected from a database, it can be administered as pre/post assessment, and with research showing its reliability and validity (delMas, R., Garfield, J., Ooms, A., & Chance, B. (2007). Assessing students’ conceptual understanding after a first course in statistics. Statistics Education Research Journal, 6(2), 28-58. <http://www.stat.auckland.ac.nz/~iase/serj/SERJ6(2)_delMas.pdf)>.  MA 103 – Quantitative Methods has a Q skill designation and primarily serves students with non-STEM (Science, Technology, Engineering, and Math) POEs. It is generally offered five times per academic year, and is the ***only Q course taken by over 15% of*** ***most graduating classes***. It, along with MA 116 – Discrete Structures, are the only 100-level courses with a Q designation, so the committee considered assessments focused on these courses. Committee member Kruse was co-author on a paper, Using Performance Tasks to Improve Quantitative Reasoning in an Introductory Mathematics Course, <http://w3.georgiasouthern.edu/ijsotl/v7n2/articles/Acc%20Art_Kruse%20&%20Drews/index.html>. This paper includes an assessment and scoring rubric for quantitative literacy. The committee elected to adopt this assessment for MA 103 and MA 116.  The introductory courses with QM designations which the committee focused on assessing are MA 130 – Calculus 1, MA 100 – PreCalculus, and EB 210 – Quantitative Business Analysis. Each of these courses is designed to meet the definitions in the handbook, which states they “must use a combination of algebraic, graphical, and numerical reasoning. Such courses should teach students how to translate problems into mathematical language, how to solve the mathematical problems, and how to interpret the solutions.” These courses serve different cohorts of students, and their associated learning outcomes, while consistent with the definition, are ultimately implemented differently. The committee met with faculty teaching these courses to develop pre/post assessments aligned with their learning outcomes. |
| **Implementation of the Q Assessment Plan, Fall 2014 Semester** |
| Assessments and Informed Consents were finalized in mid-August 2014. The response sheets for the assessments done in class included student demographic information above a “dotted line” and only the associated student id below the line, to promote anonymity while scoring.  In mid-August, before the semester began, instructors in courses were contacted to schedule times to administer the assessments, as well as to determine a count of students.  Members of the committee were scheduled to administer the pre-assessments.  Copies of the assessments and informed consents were made and distributed to committee members proctoring.  After the pre-assessments and informed consent forms were returned to Jerry, he sorted them all, confirmed that each assessment had a corresponding informed consent, and used the birthdate provided on the informed consent to mark any assessments from students less than 18 years old.  Before the end of the semester, instructors in courses were contacted to schedule times to administer the assessments, and members of the committee were scheduled to administer the post-assessments.  After the post-assessments were returned to Jerry, he sorted them, and paired each pre-assessment with a corresponding post-assessment, and confirmed they had also matched up with a valid, “18 or over” informed consent. The pre-assessments were administered before the end of the “drop/add” period, so some students took the post but not the pre, while others dropped the class. Only valid pre-post pairs were scored.  The assessments were then marked, using a random character generator, with a alpha-numeric code specific to each section and pre/post assessment, and it was written above and below the “dotted line,” so the responses could be scored anonymously. This code was used, along with the student id number, on the Q scoring rubrics as well. The alpha-numeric code made sorting the scoring sheets after the grading much easier. Student work was randomized before distribution for grading.  Kim and Matt graded the QM assessments, Jerry, Carlee, and John graded the Q assessments. Each group took care to promote inter-rater reliability. For each of the QM assessments, Kim and Matt created scoring guidelines, scored a few together, then scored the same ~10 assessments and compared results. Finally, they scored the rest separately, asking questions of the other as needed for clarification. For the Q report, Jerry, Carlee, and John graded the same set of 10 responses, then consulted on the scores, then graded another set of 10 responses and compared results. On this second set of scores, there were 5 categories, giving 50 possible score comparisons. 46 of the 50 scores had scores that differed by at most 1 between scorers, giving us confidence to proceed with individual scoring. The remaining 4 items had scores that differed by 2 (never 3). We used the following rule to record scores of these “training” protocols, “if 2 or more scores are the same, use that score, otherwise, use the mean average for the entered score.” |
| **Committee expectations of what data might show, during and after scoring sessions, January 2015** |
| The committee expected that on the whole, students would have a higher score on the post-assessments than on the pre-assessment, which is reflected in the data. However, we also had concerns regarding how seriously the students considered the post-assessment, considering that it was administered during the last week of classes and typically did not affect on their grade. There were some students in the Q and QS assessments whose post-assessment scores were substantially lower. We noted that in general the students completed their post-assessments more quickly than their pre-assessments. In particular, the CAOS reports indicate that the average time taken decreased in most sections. The proctors also noted that many of the students taking the Q post-assessment started writing their responses very quickly, before they had read all the given evidence. The committee wondered if this was simply not taking the post-assessment seriously, or some underlying issue with the perceived credibility of the evidence provided.  The committee wondered if there would be a ceiling effect on the MA 130 - Calculus 1 scores, since a large cohort (roughly 75%) of students had taken a Calculus course previously. |
| **Scoring Q pre and post responses, January 2015** |
| ***Thursday 15.Jan – Scoring Rubrics and Inter-rater Reliability 9:00am*** – Intro, solicit comments on assessment process  ***9:15am*** – Discuss “Divide and conquer” *Q assessment*– Jerry, Carlee, and John *QM assessment(s)*– Kim and Matt grading Q/QM protocols ***once?*** after  reliability testing?  ***9:30am 9:30am*** *Q assessment*– Jerry, Carlee, and John *QM assessment(s)*– Kim and Matt review and discuss performance task discuss grading criteria   (partial credit? points/question?)  ***10:00am*** ***10:00am*** *Q assessment*– Jerry, Carlee, and John *QM assessment(s)*– Kim and Matt review and discuss scoring rubric train on 4 randomly chosen responses  ***10:30am 0:30am*** *Q assessment*– Jerry, Carlee, and John *QM assessment(s)*– Kim and Matt train on 4 randomly chosen responses grade then compare 10 responses   🡪 repeat w/10 until reliability is  acceptable  ***DEMO RELIABILITY DEMO RELIABILITY***  ***Until lunch Until 2:00pm*** *Q assessment*– Jerry, Carlee, and John *QM assessment(s)*– Kim and Matt grade then compare 10 responses grade responses 🡪 repeat w/10 until reliability is  acceptable  ***11:45am*** – Lunch  ***1:30pm*** – “Where are we?”  Plan for tomorrow.  **2:00pm** – Adjourn  ***Friday 16.Jan – continued Scoring*** |
| **Committee observations during and after scoring sessions, January 2015** |
| Comments on the assessment implementation:   * How to handle students with documented learning differences with accommodations? No accommodations were given for these assessments. * Consider motivation/incentives, and ensuring that rewards were consistent (some courses counted assessments as homework). * Should GenEd requirements be considered as proficiencies or experiences? For example, should a student be required to take an upper-level GenEd course if they would be able to pass a proficiency test for an introductory course? (***see discussion below***) * The QM assessments were different, by class, while the Q and QS assessments were the same for all classes. * In the Q assessment, the category “***Evaluating evidence***” might have lower scores since this is not explicitly asked for in the prompt. * In the Q assessment grading, we decided that no documents mentioned is a “low” score, mentioning most/all documents is a “medium” score, and mentioning most/all documents + validity/relevance/credibility is a “high” score. * Information literacy is a much larger component than expected in the Q assessment, and it is a thread through the previous GenEd assessments as well. Should consider modifying the presentation in the booklets, it might help with how students perceive the credibility of the sources. * In the MA 130 – Calculus 1 QM assessment, the time interval on #3 should be increased, it shouldn’t be 1! And in the related rates problem, consider the wording “the Vol/Time is ----, the rate of increase of r is ----” indicating that the derivative must calculated at a point, it’s not a constant. * One section on MA 130 – Calculus 1 seemed to be better prepared, with higher pre-assessment scores. * It would have been helpful to have sample CAOS problems to show QS faculty. |
| **Considerations for the Faculty regarding this and future assessments of the Quantitative Skills requirement** |
| Some courses with obvious Quantitative content do not have Q designations (Math Methods in Physics, Nature of Math, etc.). While this might not have an impact on the curriculum, it is a consideration when a course’s listed skill or distribution is used to classify it for assessment. Is there a difference between a course with a significant quantitative component but without a Q/QM/QS designation, and one with a designation? The analogy would be to courses with a significant writing component but no CW designation, and courses with the CW designation.  Should we be considering assessment holistically, doing a pre-assessment on incoming freshmen and a post-assessment on graduating seniors? The focus on individual courses at least allows for some consideration of where learning goals might be met. When assessing individual courses, the committee could focus more on using embedded, pre-existing assessments.  A question from the faculty concerned whether 1-credit QS courses were sufficient?  If the GenEd requirements are proficiencies, students should be allowed to test out.  For example, if we had a second language requirement and a student came in as a freshman with complete proficiency in English and Spanish, would she still have to take two semesters of a language?  The answer to this question may help clarify the purpose of the GenEd curriculum.  If GenEd definitions are proficiencies, then yes, students could test out by demonstrating the required skill.  If GenEd definitions are processes, then no, students cannot test out, even though they can demonstrate the required skill.  The advantage of proficiencies is a clearly defined skill.  The advantage of processes is a collective experience.  The question for Juniata is how to achieve the defined skill in the midst of a meaningful collective experience.  The inconsistent student learning outcomes between courses may account for some of the differences by section of the CAOS assessment. For example, not all QS courses might devote class time to analyzing graphs. The results on some questions were better on pre than post (correct intuition getting trumped by incorrect learning), while other questions had huge improvements, which clearly indicates this is material covered in class.  The small effect sizes comparing pre and post assessment results is not unusual and mirrors national trends (statistical vs. practical significance). The committee suspects that more student learning is occurring than is reflected on the results. One possible explanation is that the scoring was too “coarse,” only one or two points possible per part could dampen gains. Another explanation could be that the learning outcomes chosen when designing the assessments don’t mirror the actual learning outcomes. |

Would it be possible to design new assessments for each course, could we design one for each that would result in larger gains? The Calc assessment showed the most gains, since it had difficult question (Related Rates)?