Annual Assessment Report to the College 2011-12

College: _ Science & Mathematics _____

Department: _ Geological Sciences _____

Program: ___B.S._____

Note: Please submit report to your department chair or program coordinator, the assessment office and to the Associate Dean of your College by September 28, 2012. You may submit a separate report for each program which conducted assessment activities.

Liaison: __Matthew d'Alessio _____

1. **Overview of Annual Assessment Project(s) (optional)**

1a. **Assessment Process Overview:** Provide a brief overview of the assessment plan and process this year.

Highlights include:

1. Revisit SLO's for B.S. degree. (Result: reworded and consolidated them)
2. Discuss and construct SLO's for M.S. degree (Result: New SLO's for that program)
3. Created drafts of assessment rubrics for SLO 4 (Oral and written communication)
4. Collected data in gateway and capstone courses for SLO 1 (Conceptual understanding)
5. Analyzed data from SLO 1. (Result: revisions to undergraduate curriculum at the course level in key areas; revisions to assessment instrument to clarify questions)
6. Collected data at the First Annual Geologic Problem Solving Night to collect demographic data about our students and to assess SLO 2 and part of SLO 3. A huge success!
7. Preliminary analysis and discussion as a faculty of student submissions from the Problem Solving Night (Result: Awareness of current deficiencies; philosophical discussion of solutions; agreement that revisions to our undergraduate program should address these deficiencies.

2. Student Learning Outcome Assessment Project: Answer questions according to the individual SLO assessed this year. If you assessed an additional SLO, report in the next chart below.

2a. Which Student Learning Outcome was measured this year?

1) demonstrate conceptual understanding of different earth materials and the processes that shape them throughout their history;

2b. Does this learning outcome align with one of the following University Fundamental Learning

Competencies? (check any which apply)

NO. This is discipline-specific content knowledge.

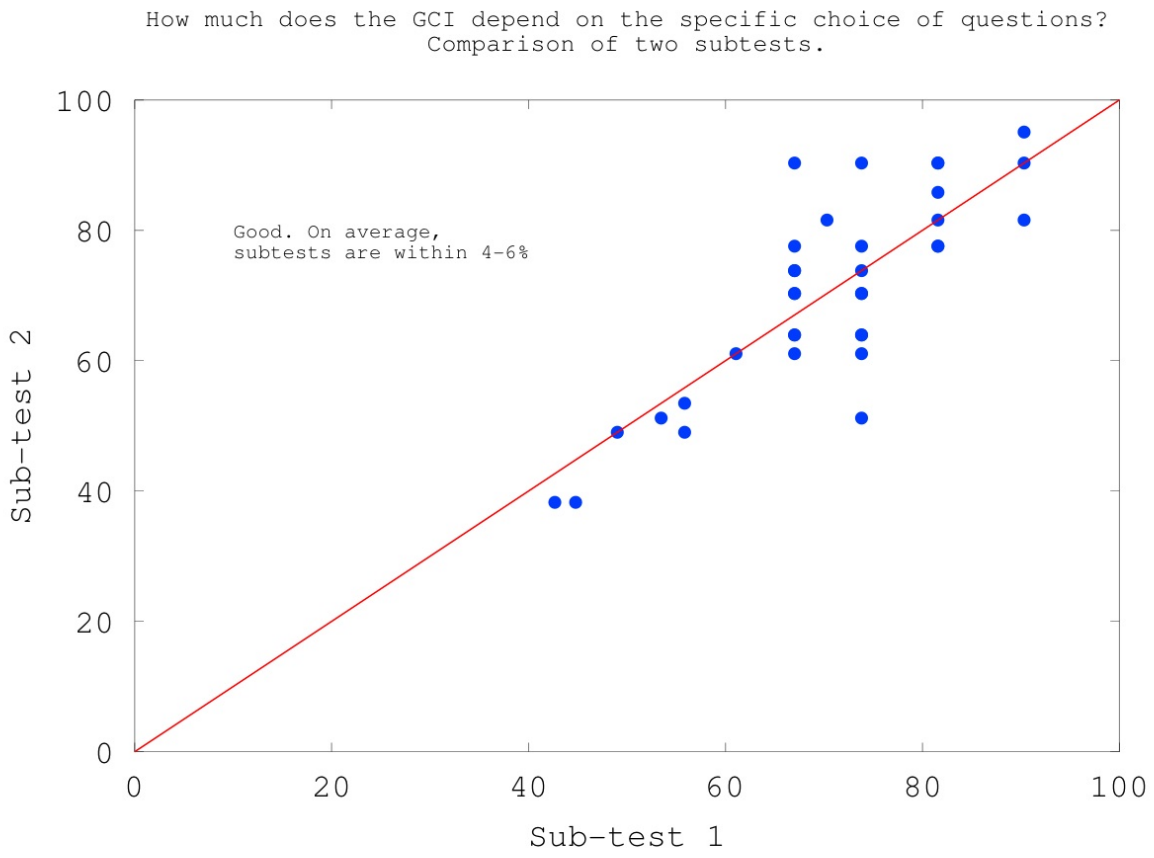
2c. What direct and indirect instrument(s) were used to measure this SLO?

We developed our own version of the Geoscience Concept Inventory, a national test designed to assess conceptual understanding of some big-picture ideas in Earth Science. Questions on the GCI are well validated.

A GCI "test" consists of 15 sub-items from a much larger pool of validated questions, with questions selected based on applicability to the individual course or program being assessed.

As a department, we ranked all 80+ questions in the GCI pool based on how important it was to ask. About 50% of the departments' faculty contributed to the ranking process and two questions were actually unanimously chosen as best. As the assessment liaison, I then created a complete test that satisfied the requirements of the GCI for breadth (and made sure to include the few required questions so that we can place our results in the national context). Faculty then reviewed the completed assessment and made comments on individual questions. We modified the wording compared to the validated questions for about 5 items so that they better fit the language we use in our own department for the concepts.

Our version of the GCI had 24 questions so that we could calculate two almost-independent GCI scores for each individual. We find good correlation between the two sub-scores for high performing and low performing students. Mid-range students have selective knowledge, so it is not surprising that their score depends on the choice of questions. Median repeatability is 4% between sub-tests, but scores differed by 20% for a few students.

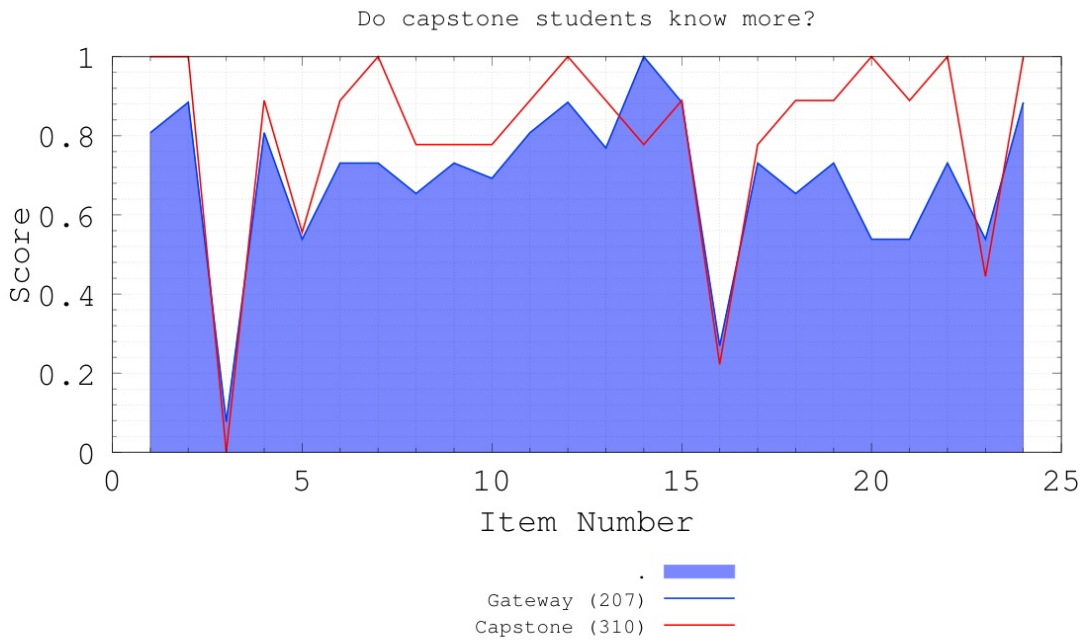


2d. Describe the assessment design methodology: For example, was this SLO assessed longitudinally (same students at different points) or was a cross-sectional comparison used (comparing freshmen with seniors)? If so, describe the assessment points used.

This is the first year of our GCI use, so we performed a cross-sectional comparison between students in our gateway course for majors (GEOL 207, n=28) and an upper-division required course that is typically the last required course before students take electives (GEOL 310, n=9). We do not have a true capstone course in our major, but we refer to GEOL 310 as our capstone for the purposes of this report.

2e. Assessment Results & Analysis of this SLO: Provide a summary of how the evidence was analyzed and highlight important findings from collected evidence.

Below is a graph comparing the results on the 24 individual items in our gateway and "capstone" course.



We focused on 5 key questions that had exceptionally low scores and showed no major difference between gateway and capstone. Our faculty examined each of those questions:

1. Two questions are considered "trick questions" that were poorly worded or didn't test what we wanted them to test.
2. One question was about Earth's climate and is a required question on the national implementation of the GCI. Our department focuses on solid-Earth processes, not atmospheric and climate processes. The low score on this fundamental Earth science concept forced us to admit this limitation in our department.
3. We regarded two questions as serious failure by us/our students. Both were deemed fair questions about a fundamental Earth process, but students scored low on them. On one question, there was some improvement from gateway to capstone. On the other, there was no improvement. Either way, we felt that 100% of capstone students should have answered this question correctly. See below for discussion.

Benchmarks:
Our stated goal was that 75% of our graduates would get more than 75% correct on the GCI.

In our gateway course, 46% of the students met the 75% benchmark.*
In our capstone course, 78% of the students met the 75% benchmark.*

*(this analysis excludes the two questions that the faculty deemed 'invalid' despite the fact that they have undergone extensive validation nationally. If we include those questions, only 33% of our students met the 75% benchmark (meaning that a large portion of students score between 70-75% on the uncorrected scale).

2f. Use of Assessment Results of this SLO: Were assessment results from previous years or from this year used to make program changes in this reporting year?
Type of change:
changes to course content/topics covered_____

Ring of Fire question: In our classes, we focus on plate boundaries (which often but not always exist at the edges of continents). During all this discussion, we offer much less discussion of passive margins (continental coastlines that are not active plate boundaries. We suggested that we should include additional discussion of passive margin processes.

Continent v. Ocean rock age question: Introductory classes introduce plate tectonics by citing the historical progression of discovery. Early researchers on plate tectonics used symmetric patterns of magnetic striping, but understanding how this data set works is extremely complicated – it requires understanding how Earth's magnetic field works, how rocks record the ambient magnetic field, and familiarity with bizarre magnetic field reversals. A much simpler data set exists in the form of sea-floor data. In introductory classes, we suggested that we should focus on the age data when teaching about plate tectonics. This is hard because textbooks use the complicated data set (and spend several pages explaining it!).

assessment instruments:
We will revise the questions that are considered "trick questions."

2a. Which Student Learning Outcome was measured this year?

2) **identify geologic problems and develop testable hypotheses that would aid in their solution both independently** and in collaboration with others

3) **demonstrate skills in** standard data-gathering and **data-analysis method**s in both lab and field settings

2b. Does this learning outcome align with one of the following University Fundamental Learning Competencies? (check any which apply)
Critical Thinking_____XX (SLO 2)_____
Quantitative Literacy_____XX (SLO 3)_____

2c. What direct and indirect instrument(s) were used to measure this SLO?

We utilized a "progressively revealed problem-solving assessment." It is based on some of the case study assessments done by the Biology department at University of New Mexico (references available). The analysis tests 5 independent skills in scientific problem solving:

1. Hypothesis generation
2. Planning investigations

3. Analyzing data (*quantitative literacy*)
4. Evaluate Conclusions (*critical thinking*)
5. Reflect on Conclusions and Formulate Next Steps

We present students with a single scenario ("The Case of the Muddy Fish" for this year's prompt). They respond to the 5 key skills listed above in a series of 5 prompts. With each new prompt, new information is revealed. While all building on a single scenario, the information required for each skill is provided within each skill's prompt. In other words, students are not penalized for "not-knowing how to answer a previous question" because a valid correct answer to the previous question is revealed each time the student moves forward. The situation is novel enough that no students have prior knowledge tackling this specific problem. Though they certainly do draw on some background knowledge in geology, we are hoping to test these independent scientific process skills. We implemented the exercise electronically in Moodle. Here is how we implemented this system:

1. After an introduction to the situation, students must type in up to 5 hypotheses about what has caused the observed increase in mud in the river.
2. Upon submitting that response, they are asked to pursue one specific hypothesis chosen by the exam writers (it may not necessarily appear in the list submitted by the student). To test their ability to plan the investigation, they enter in a list of different types of relevant data that would help them test the given hypothesis.
3. After hitting submit, they are presented with actual data chosen by the exam writers (again, it may not coincide with any of the data they requested in the previous step). Students are asked to narrate an analysis of the graphs, looking for trends and relationships. They must also decide whether or not their original hypothesis is supported by the data set.
4. After hitting submit, they are asked to defend a specific conclusion chosen by the exam writers. They enter in a response to an fictitious anonymous peer reviewer defending their position and/or acknowledging the viewpoints in the review comments.
5. After hitting submit, they are asked to reflect on the entirety of the scenario thus far. What new questions do they have that they would like to pursue based on the data and analysis they've seen.

Responses to each prompt are relatively short paragraphs or bullet points.

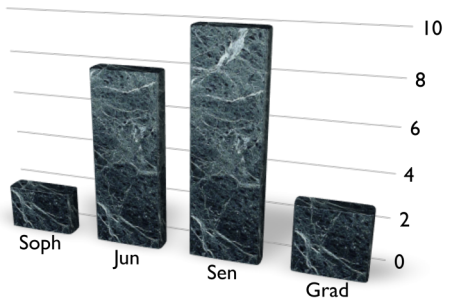
2d. Describe the assessment design methodology: For example, was this SLO assessed longitudinally (same students at different points) or was a cross-sectional comparison used (comparing freshmen with seniors)? If so, describe the assessment points used.

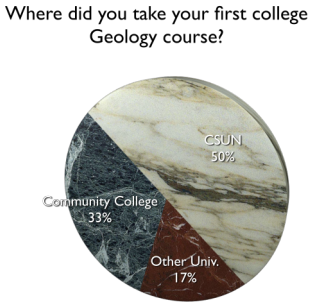
Lacking a formal capstone course in our major, we experimented with a new method for obtaining assessment data. We sponsored the "First Annual Geologic Problem Solving Night." We invited all our undegraduate majors to attend an evening of free pizza and ice cream. We asked faculty NOT to offer extra credit for participation. Instead, we asked faculty in our core undergraduate courses to provide one question from their upcoming final exam. We would reveal the question to students and discuss the answer. This was a STRONG motivator to students, and we even had graduate students enrolled in these core classes ask if they could attend, too.

Students sat at laptops and worked individually on the progressively-revealed problem solving exercise on Moodle. Considering this was an evening activity the week before final exams, we were thrilled by the results! Twenty-three students volunteered to show up and take the assessment. While working through the problems, students were absolutely silent and focused. The quality and seriousness of their responses are a testament to this focus. We feel that this experience is a strong counter-example to the

idea that students need to be graded on assessment items in order for them to take them seriously.

From the perspective of rigor, our sample is highly biased. The participants were self-selected. We can state that they come from a wide range of backgrounds and class-levels. (We collected overview demographic data, including their history in our department and pathway to the geology department through community colleges and earlier experiences).





2e. Assessment Results & Analysis of this SLO: Provide a summary of how the evidence was analyzed and highlight important findings from collected evidence.

In the future, we will develop a robust rubric for each skill subscale and a specific implementation of that rubric for each year's case. However, we had faculty score the responses "holistically" for our preliminary analysis of the pilot data. We presented them with a definition of the skill and a copy of the student prompt. Before they were allowed to view student responses, we asked them to formulate their own response and think about what they expected of undergraduates in our department. After discussing this holistic 'rubric' with another faculty member, they were allowed to evaluate student responses, adjusting as necessary after seeing a sample of student responses. Each response was scored using some variant of a three point scale:

1. Above expectations for an undergraduate about to receive a B.S. in our department
2. Met expectations
3. Below expectations

At this point, we have only analyzed the entire data set as a single unit. We have not analyzed responses of sub-groups such as sophomores v. seniors). We will perform that analysis during this academic year, but it may have limited meaning due to the small numbers. We look forward to pooling results from this coming year's responses to previous years in order to perform a more robust cross-sectional analysis. Since individual responses are stored by Moodle ID, we can also track longitudinal changes. We hope these will be robust in aggregate, but there may be concerns because prior knowledge in the specific sub-discipline of the case study does impact a students' individual performance (the tests are designed to avoid this, but it's impossible to completely disaggregate process skills from background knowledge – see the extensive literature on the definition of "expert" thinking).

Results for all 5 skills are below:

	Below our expectations	Met Expectation	Above Expectation
Hypothesis	5	9	8
Planning Investigations	4	12	6
Analyzing Data	15	6	1
Evaluating Conclusions	12	4	Maybe 4
Reflecting	5	10	4

When looking at the data table above, faculty remarked that students clearly lacked skills in analyzing data and evaluating conclusions.

Benchmarks:
In the past, we had not set a benchmark for this SLO or any of these specific skills. Overall, the faculty seemed satisfied with the results for hypothesis generation (78% met or exceeded expectations) but seemed displeased by the Analyzing data (32% met or exceeded expectations).

2f. Use of Assessment Results of this SLO: Were assessment results from previous years or from this year used to make program changes in this reporting year?
Type of change:

changes to course content/topics covered_____

We would like to emphasize how these data were analyzed: The entire faculty reviewed the all the responses about at least one skill. This reflects substantial buy-in by the faculty to this process. They were genuinely interested in contributing to the process and genuinely intrigued as the results unfolded before us.

Perhaps the biggest product of this process has been a clearer articulation of the specific sub-skills that make up the scientific process. As we redesign our curriculum, we are going to strive to explicitly address each of these skills.

Specifically, we discussed approaches to better teach data analysis and practice evaluating conclusions. Here are some specific ideas we discussed:

- Too often, feedback about student work is too far removed from their completion of projects. We provide extensive written comments on projects and reports, but provide few avenues for students to act on those comments. Without much motivation to read the comments, they never internalize the feedback and use it to improve their data analysis skills. Greater use of peer assessment of work

can help provide real-time feedback.

- Instructors can also do a better job modeling their own data analysis approach. Through think-aloud exercises, they can reveal how they would approach the problem. Explicitly narrating the process can sometimes help students construct their own schemas for problem-solving. The process also makes the schema more explicit to the instructor so that he/she can do a better job recognizing where students might be missing key steps.

Implementing these changes is an ongoing process by individual faculty in their courses.

assessment instruments:
We actually tried two progressively revealed problem solving cases. We will spend this academic year looking at the second case to see where it worked well and where it didn't. By comparing the two cases, we can hopefully find some common elements of success!

Some programs assess multiple SLOs each year. If your program assessed an additional SLO, report the process for that individual SLO below. If you need additional SLO charts, please cut & paste the empty chart as many times as needed. If you did NOT assess another SLO, skip this section.

3. **How do your assessment activities connect with your program's strategic plan and/or 5-yr assessment plan?**

We are making good progress on our assessment plan. The curriculum committee is currently busy designing a new sequence for the B.S. program, as planned. Our discussions focus around questions of how to make authentic assessment pervasive throughout the curriculum, not just for the purpose of University assessment.

We were probably over-ambitious to state that we would perform so many embedded assessments of PLO 1 as described in our 5-year-plan. In future years, this might be possible. However, our current activities included so much planning and design of activities and rubrics that we did collect that we did not leave enough time for these data collection activities. However, we did complete some data collection on two SLO's and even with just these data, we were able to have substantive discussions about our teaching practices! It makes us wonder if our five year plan should have substantially less data collection and substantially more time included for discussion of the results in more limited efforts.

4. **Other information, assessment or reflective activities or processes not captured above.**

5. **Has someone in your program completed, submitted or published a manuscript which uses or describes assessment activities in your program? Please provide citation or discuss.**