Take the Next Step in Program Learning Assessment: Collect & Review Evidence of Learning
Facilitator: Monica Stitt-Bergh

Bring your program student learning outcomes and leave with a plan to assess how well students are meeting those outcomes! The facilitator will share good practices for collecting and reviewing evidence of student learning, tips, and examples. During the workshop, attendees will have the opportunity to consider various options and discuss which might work best for their program given its size, resources, and assessment goals.

By the end of the workshop, attendees will
Be able to list different ways to collect and review evidence of learning
Be able to describe benefits and drawbacks for the different ways
Have at least one meaningful and manageable plan for their own program

Who should attend: faculty and staff
Level: Beginner
Format: Lecture + activity

Date/time/location:
Thursday, February 27, 2014, 2:30-3:45 pm, KUY 106
OR
Monday, March 3, 2014, 10:30-11:45am, KUY 106
Yao and I are the Assessment Office. We were hired to help programs improve student learning through program assessment. Workshops are one method we use. You can also contact us for individualized consultations or visit our website for helpful materials.

By the end of today’s session, you will be able to
a) list different ways to collect and review evidence of learning and
b) describe benefits and drawbacks for the different ways.
And finally, you’ll have drafted a
c) a meaningful and manageable plan for your program.
To make sure we are on the same page, here is our definition of program assessment.
I’ll point out that our office deals with student learning at the program level. Others on campus deal with other aspects such as graduation and retention rates, student satisfaction, etc.

This is the standard program assessment cycle that faculty undertake. As the definition stated, first, faculty make the learning outcomes explicit. Then, they offer students opportunities to learn. Next they collect evidence of learning and evaluate the evidence. The final and most important step is using the results for improvement or program evolution to address new needs and demands.
Today’s focus is on the collection and evaluation of evidence of learning.

I have two brief examples to share with you.
First, let me tell you what the English (BA) program is doing this year. The assessment coordinator consulted the curriculum map and found two program outcomes that students have opportunities to learn in four required courses in 100-300 level courses: 1) critically interpret texts and 2) use secondary sources. These two outcomes are assessed in their 400-level “Studies in English” course series.
The assessment coordinator selected the 400-level, “Studies in English” course series as the ideal place to collect evidence. Students were already completing a major research paper in the course series. The course series is taught by 6-8 different faculty each semester and each course in the series has a different subject, for example, in spring, one course studies Asia-Pacific Literature and another studies Fairy Tales & Social Struggles.

The assessment coordinator gathered together the faculty teaching the course and they agreed on common elements for the final research paper guidelines as a starting point. So even though students write on different topics, the genre and basic elements of the paper will be the same: in this case, the research paper will be an argument that uses outside sources to support and extend the student’s idea; students must critically analyze and interpret at least one text; the paper is 10-15 pages and follows MLA Style Guide.
Faculty teaching the senior, required course:

a) Jointly create a rubric to evaluate the final papers
b) Evaluate seniors’ papers using the rubric

The faculty also gathered together to agree on a rubric that will be used to evaluate the final research papers. They started with the learning outcomes and two existing rubrics. They modified the rubrics to create one rubric. Faculty will use the rubric, but they can also add course-specific criteria if they want to.

Similar to the English BA program, faculty in CEE used the curriculum map as a tool to map out which courses would address the outcomes and locate where evidence could be collected. Students in CEE practice design and communicating in 11 required courses.
The faculty agreed to collect evidence from four 400-level courses. The faculty who teach these four courses evaluate student exams and assignments that are aligned with the two outcomes and they provide the assessment coordinator with a summary “performance appraisal” report based on student performance.

These two programs found that they could use existing assignments and exams with a few modifications to ensure the assignments would provide information about the outcomes under investigation. I chose to share these two examples because using existing student work and rubrics are an excellent way to make assessment meaningful to faculty and students and to make the process feasible.
There are many different types of evidence of learning in addition to the ones I just mentioned, research paper, course assignments, course exams.

Many possibilities!
“DIRECT” AND “INDIRECT” EVIDENCE

A) Write down the name of each state’s capital city:
   Alabama: ________. Alaska: _________ . . . .

B) How many state capital cities can you name?
   □ 0      □ 1-20     □ 21-40     □ 41-49     □ 50

C) Write a research paper in which you . . .

In the field of assessment and evaluation, there are two handy terms that help categorize the type of evidence collected from students: direct evidence and indirect evidence of student learning. Here are three items. Write down on scratch paper which you think give direct evidence and which give indirect evidence of student learning.
A) Write down the name of each state's capital city:
Alabama: ________. Alaska: ________ . . .

B) How many state capital cities can you name?
- 0
- 1-20
- 21-40
- 41-49
- 50

C) Write a research paper in which you . . .

A and C produce direct evidence of student learning. B gives indirect evidence of student learning: from B, we will find out whether students believe they know the capital cities while item A will tell us exactly the capital cities that students know and don’t know.

For learning outcomes assessment, faculty should be gathering direct evidence of student learning. Please do not rely solely on indirect evidence because that is based on student perception, not expert faculty evaluation of student knowledge or skill. Indirect evidence has it’s place and uses, but it does not tell us enough about what students know and can do.
On the handout “Match the Method,” I’ve listed the most common types of evidence used by our undergraduate programs. Take a few minutes to complete the second and third columns. I will not collect these; they are for your eyes only.

There are only two indirect methods listed: student interview and survey.
On your handout, you can use the space in the last two columns to take notes.
The answers to the “match the method” questions are on the slide. E.g., Description “D” corresponds to Capstone paper.

**Capstone, senior thesis, final exhibit (culminating project)**
Perhaps too many students leave college feeling like they took a bunch of interesting classes, but they didn’t have a chance to tie everything together. A capstone project can help students put all the pieces together. At the graduate level, the thesis, seminar paper, and dissertation are graduate students’ capstone project. For these reasons, capstones can benefit students and faculty.

On campus, capstones are scattered across campus. Honors students complete a senior thesis. Engineering has a senior design project course in which students are guided through an engineering project from start to finish. Capstones have great benefits for students: they see their accomplishment as they draw on skills and knowledge from previous courses to complete this major culminating project.

Strength=students are highly motivated and we know from research studies that highly motivated students score higher on exams.
Limitation: Faculty need to setup the project so it’s meaningful and doable. They need time needed to plan, organize, and administer so that regardless who teaches the capstone, the format is similar.

**Course assignments**
An example of using what’s called “embedded assessment,” is to use course assignments. For example, if you want to know how well students are meeting an outcome about problem solving, collect problem solving assignments that students already do as part of their coursework. Evaluate and discuss a sample of them in a 3-4 hour faculty meeting.
Students are motivated because it counts toward their course grade. Some faculty fear it impinges on their ability to teach what they want.

**Course test**
Another example of “embedded assessment” is using a multiple-choice test or essay exam that students take as part of regular course requirements. For example, a subset of questions can be directly aligned with a program.
outcome. For example, language courses agreed to include a handful of the same questions on the 101, 201 and 202 final exams. They then compared the results to determine student development over time. Students are motivated because it counts toward course grade. On the downside, some faculty fear it impinges on their ability to test what they want.

**Observation**

Some outcomes don’t lend themselves to written documentation. Performance fields such as Art & Dance are obvious. But Science departments may want to observe students do lab experiments; Education observes student teaching performance; Business may observe group leadership skills. The key is to have an “observation checklist” or scoring guide so observations are done fairly across students and observers. This may take some training. For Speech courses, the teaching assistants practice using a rubric by watching a video-recorded student speech, evaluating using a scoring guide, and then comparing their scores & discussing to reach consensus.
Portfolio
One of the most robust methods is the portfolio. It typically allows many outcomes to be evaluated. It can show growth and changes (a “developmental portfolio”) or show accomplishment (a “showcase portfolio”). Because it involves multiple pieces of student work and typically a reflection piece, it’s comprehensive and research has shown that asking students to gather pieces and reflect increases their metacognitive skills. In addition, students have pieces to show employers and graduate schools. For these reasons, it’s beneficial to students and allows faculty a multi-faceted look at student performance.

We have examples of portfolios on campus. Art has been using portfolios for years: students exhibit their artistic pieces in a portfolio. Dance department requires a portfolio: all seniors submit a photographs, dance reviews, scholarly papers, a reflective self assessment, and a resume.

It’s strength is that it’s comprehensive and useful to students and faculty. Unfortunately, you get what you pay for: it’s a great way to collect data, but it is time consuming to plan, organize, administer, and evaluate. I recommend that programs start small, with embedded assessment and work up to a portfolio. I don’t encourage programs to start with the portfolio.

Pre-post test
If your program is interested in change, administer the same or similar test at the beginning and end of a program. Speech and Human Nutrition programs had students take a content test while they were early in the program (100- and 200-level) and then take the same test at the end of their degree as seniors.

While this is basically the only method that allows conclusions about student growth and development, there are limitations and challenges: it’s difficult to create a good pre-post-test, knowledge of advanced statistics is often necessary because of student attrition and missing data, and there can be a pre-test effect such as just by virtue of taking the pre-test the students will perform better on the post-test.

Standardized Test
Programs using standardized exams either purchase the exam or their students pay for the exam themselves because it’s the professional licensing or certification exam. The main benefit is national or cross-group comparisons can be made. The main limitation is that motivation is often low for the students unless it’s a licensure exam that they really want to pass to become a practicing professional.

In addition, motivation is often low for the faculty too, unless their accrediting agency bases accreditation on a certain pass rate (such as in the case of the law school).

<table>
<thead>
<tr>
<th>Evidence of Learning</th>
<th>Strength</th>
<th>Limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Portfolio [B]</td>
<td>Comprehensive; Reflective</td>
<td>Resource intensive</td>
</tr>
<tr>
<td>Pre/Post Test [C]</td>
<td>Demonstrate growth</td>
<td>Pre-test effects</td>
</tr>
<tr>
<td>Standardized Test [A]</td>
<td>National comparisons</td>
<td>Low motivation unless licensure exam</td>
</tr>
<tr>
<td>Supervisor/ Employer Evaluation [H]</td>
<td>External feedback</td>
<td>Collection of evaluations</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Lack of standardization across evaluations</td>
</tr>
</tbody>
</table>
Supervisor/Employer Evaluation
I see supervisor evaluations most frequently in programs that require a practical experience such as nursing, social work, communication sciences & disorders. Some programs such as kinesiology and rehabilitation sciences ask employers to rate alumni on the learning outcomes. Having external feedback is an excellent quality check for the program and it helps programs prepare students for post-school experiences.
Unfortunately, supervisors and employers are busy people too and collecting the evaluations takes someone who can make multiple contacts and make the filling out of the evaluation form easy.
Because supervisors and employers are in different locations, different work environments, it’s difficult to have any standardization across evaluations—one person’s excellent may not be another’s, and so forth.
I’m going to switch gears now and talk about how to evaluate evidence. There’s three main methods:

- A rubric or simple checklist
- Observation checklist
- Answer key for exam

A tip: start with an existing rubric or observation checklist and modify it as needed to fit your outcomes.
The purpose of assessment is program improvement or program evolution so that students achieve the desired learning outcomes. For assessment to work, it needs to be meaningful and viable. So, when deciding from where to collect and what to collect and how to evaluate, here’s a few factors to keep in mind.

First, the evidence of learning must be considered credible by the program faculty.

Second, it must be feasible to collect and evaluate once a year or once every other year.

Finally, the collection and evaluation process needs to result in useful information. I recommend creating hypothetical results and seeing if faculty can use them to guide program decision making.

On the back of the “Examples” handout is a simple bar chart which is sufficient for presenting results. The results are ordered by most achieved to least achieved. In this case, students did less well on “creativity” and “simplicity” in their design project. Given a program’s assessment plan, I will create a simple chart such as this and discuss it with faculty BEFORE data collection. The discussion can reveal whether the assessment plan needs to change and if faculty are willing to make changes based on results.
Before you start working on your own plan, I will point out a few highlights from the “Examples” handout. Given the recent Winter Olympics, I took a Bronze, Silver, Gold approach.

I underlined key phrases in each column. The bronze example made it to third place because the evidence is a pre-post-test within a single course, graded by only the course professor.

The silver example is better than bronze. It’s similar to the English and Civil Engineering examples I provided earlier. In the silver column, you see that they used the curriculum map; the evidence is an extended project which has agreed-upon requirements. Three professors are evaluating student work after collaborating and consulting together. They jointly created a rubric. Their findings from different courses are combined into a single set of results which is an indication they are taking the program-view (not the individual course view).

The gold example is in first because two faculty members independently evaluated the student work, not just one person as is seen in the other examples. Because this requires more work for faculty, you’ll see that in the logistics row, they evaluated a random sample of student work to keep the work load manageable.
YOUR TURN

- Create an assessment plan (*Create a Plan worksheet*)
- Pair & Share
Mahalo! I appreciate your time and I’ll answer questions.
### Evidence of Student Learning Workshop
#### Match the Evidence/Method With the Correct Description

<table>
<thead>
<tr>
<th>Description</th>
<th>Evidence/Method</th>
<th>Direct or Indirect?</th>
<th>Strengths</th>
<th>Limitations</th>
</tr>
</thead>
<tbody>
<tr>
<td>A. Students take a test that is administered, scored, and interpreted in a standard manner; often commercially developed and administered to many (e.g., students across the nation).</td>
<td>Capstone Paper, Senior Thesis, Final Exhibit</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>B. Students gather a collection of their work to demonstrate growth and/or showcase achievement; often includes students’ reflections on the collection.</td>
<td>Course Assignments (“Embedded Assessment”)</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>C. Students take a test administered at the beginning and again at the end of a program in order to determine change or growth.</td>
<td>Course Test (“Embedded Assessment”)</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>D. In their senior year, students integrate the knowledge and skills they have acquired throughout the program to produce work that demonstrates achievement.</td>
<td>Observation of Student Performance/Activity</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>E. Students/alumni report their beliefs about their knowledge and/or abilities and opinions on aspects of their educational experience.</td>
<td>Portfolio of Student Work</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>F. Students’ performance of a task or activity is rated/scored by an unobtrusive observer, often using an “observation checklist.”</td>
<td>Pre-test/Post-test</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>G. Students take tests/exams as part of a course that include questions directly related to program SLOs. The test/exam results are used for course grades and program assessment.</td>
<td>Standardized Test</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>H. Typically a written evaluation of students/alumni performance in a work setting.</td>
<td>Student Interview</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>I. One-on-one dialog with a student to determine his/her perception regarding learning outcome achievement, contributing academic experiences, and application to future plans.</td>
<td>Supervisor Evaluation of Intern or Employer Evaluation of Alumni</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
<tr>
<td>J. Students complete course assignments that are directly related to program SLOs; the assignments are used for course grades and program assessment.</td>
<td>Survey: Student or Alumni Survey</td>
<td>D I</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Evidence of Learning Workshop: **Examples of Data Collection and Evaluation**
(undergraduate degree program)

<table>
<thead>
<tr>
<th>Assessment Activity</th>
<th>Bronze Gender Studies*</th>
<th>Silver Neuroscience*</th>
<th>Gold Engineering*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Type &amp; Location of Evidence</td>
<td>The assessment coordinator and a course professor agree to use the professor’s <strong>pre- and post-test in a senior course</strong>.</td>
<td>An assessment committee <strong>uses the program’s curriculum map</strong> to identify three senior courses that significantly address the outcomes of interest. These courses require an <strong>extended design project</strong>. The professors agree to a <strong>shared set of project requirements</strong>.</td>
<td>The assessment committee and the department faculty <strong>consult the curriculum map</strong> and agree to use the capstone course and the semester-long <strong>culminating project</strong> that students complete as part of the capstone course. The professors agree to a <strong>shared set of project requirements</strong>.</td>
</tr>
</tbody>
</table>

| Person(s) Evaluating the Evidence      | A **single person**—the professor of the course. | **Three professors**. Each professor teaching an identified course evaluates their students’ design projects **after collaboration and consultation about performance expectations and how they will apply the scoring rubric**. Ideally, this includes the faculty reviewing and discussing 2-3 pieces of student work to reach a shared understanding. | **At least two faculty independently evaluate** each project **after participating in a discussion and review of at least 3 student projects to create a set of shared expectations and similar understanding of “what’s good.”** |

| Tool to Evaluate the Evidence          | Exam scoring key created by the professor of the course. | Rubric. The professors and the assessment committee **jointly modify an existing rubric**. | Rubric. The course professors, assessment committee, and other interested parties (e.g., students) **jointly modify an existing rubric**. |

| Logistics and Report of Results        | The professor of the course submits aggregate results to the assessment committee. 100% of the students are included in the summary. The assessment coordinator reports the findings to the department for action. | The professors submit aggregate results to the assessment committee. 100% of the students are included in the summary. The assessment coordinator consolidates the three sets of results into a **single set** and the committee reports the findings to the department for action. | The professors teaching the capstone courses collect the students’ final projects and give a **random sample** of redacted copies to the assessment committee. Each student project is **independently scored by two faculty members**. Their two scores are combined so each project has a single total score. The **results are tallied and aggregated**. The assessment committee creates a final report and presents to the department for action. |

* Pseudonym

continued on back
Design Project Assessment Results (spring 2014)

[Note: when presenting/distributing results, include the rubric; post examples of student work on a secure site]
Evidence of Learning Workshop: **Create a Plan to Collect and Evaluate**

1. Choose 1 or more of your program student learning outcomes (SLOs).
   
   *Tip: choose an outcome(s) that faculty really care about right now*
   
   Outcome(s) selected =

2. Choose an assessment question (or write your own):
   
   - [ ] What percent of students near graduation have met our expectations for the SLO(s) listed above?
   - [ ] To what extent are students near graduation achieving the SLO(s) listed above?
   - [ ] ____________________________

3. What type of direct evidence will answer the assessment question, is credible to faculty, and feasible to collect and evaluate? [select more than one if desired]
   
   - [ ] Capstone paper, senior thesis, final exhibit: ____________________________
     (e.g., senior art exhibit, single-author research paper)
   - [ ] Course assignment(s): ____________________________ (e.g., research paper)
   - [ ] Course exam, test, quiz
   - [ ] Observation of performance: ____________________________ (e.g., oral presentation, dance recital, piano recital, proposal oral defense)
   - [ ] Portfolio of student work
   - [ ] Pre- and post-test
   - [ ] Standardized exam
   - [ ] Supervisor/employer evaluation of performance
   - [ ] Other. Please name ____________________________

4. Does the evidence already exist—are students already performing, producing work, or taking a test aligned with the outcome(s)?
   
   *Tip: consult the curriculum map and contact professors teaching the appropriate courses*
   
   - [ ] Yes. Please note where: ____________________________
     (e.g., in ENG 489 and 490, as part of clinical practice)
   - [ ] No, we will need to take steps to ensure students perform, produce the work, or take a test.
     Briefly describe possible steps:

     ______________________________________________________________

     ______________________________________________________________
5. Who will evaluate the evidence? [Note: this may not apply to standardized exams which are typically evaluated by the organization that administers the exam]

6. How will the evidence be evaluated?
   - Rubric
   - Exam answer key
   - Observation checklist
   - Other: ____________________________________________________________

7. Create a timeline and note the lead person for each activity.

<table>
<thead>
<tr>
<th>Month, Year and Lead Person</th>
<th>ACTION</th>
<th>Estimated time (Example)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Lead a collaborative session to select outcome(s) and form an assessment question</td>
<td>1 hour</td>
</tr>
<tr>
<td></td>
<td>Communicate selected outcome(s) and assessment question to faculty in department</td>
<td>30 minutes</td>
</tr>
<tr>
<td></td>
<td>Lead the work with appropriate faculty to prepare for the collection and evaluation of evidence</td>
<td>3 hours</td>
</tr>
<tr>
<td></td>
<td>Collect and evaluate evidence</td>
<td>8 hours</td>
</tr>
<tr>
<td></td>
<td>If applicable:</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Who will lead the collection and storage of the evidence?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Who will coordinate and lead the faculty session to evaluate the evidence?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>- Who will aggregate/consolidate results into a single report?</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Create a report of results</td>
<td>2 hours</td>
</tr>
<tr>
<td></td>
<td>Present and discuss results with faculty in department to create a plan to use results</td>
<td>1 hour</td>
</tr>
<tr>
<td></td>
<td>Draft a plan to use results</td>
<td>2 hours</td>
</tr>
</tbody>
</table>

TIP: Collect & evaluate evidence from a random sample of students.