A utilization-focused approach to developing, mapping, and assessing student learning outcomes in college foreign language programs

©Yukiko Watanabe, John Davis, and John Norris
Foreign Language Program Evaluation Project
National Foreign Language Resource Center,
University of Hawaii at Manoa, 2012
Module IV
Assessing Outcomes and Utilizing Findings

- Developing SLOs
- Mapping SLOs
- Situating
- Utilizing
- Reflecting
- Planning
- Implementing
About Module IV

Stating and mapping outcomes (see Modules II and III) are the first steps toward creating shared educational expectations (and holding your program accountable). They lead to a public statement about what learners should expect, what they will learn, what they will experience to reach the outcomes, and ultimately why that is all of value. However, while discussing, clarifying, and mapping outcomes can build consensus among those involved in the process and may lead to curriculum development, stating and mapping practices themselves are not enough to make any claims about student achievement and program effectiveness. Module IV walks you through outcomes **assessment**, a systematic and reflective type of inquiry for examining and understanding student achievement to improve educational effectiveness.
Outcomes assessment is a **means for changing and improving student learning, not an end**. It enables the program to modify and adjust teaching practices—on an evidentiary basis—by gathering information on the extent to which learners achieve the outcomes the program expects. In fact, accreditation agencies do not look at how much learners are achieving program goals; rather they are interested in whether programs are monitoring program achievement and using outcomes assessment findings to make improvements in education.
How to use the module

Because outcomes assessment should be a collective endeavor, we encourage you to go through this module together with your colleagues. The module includes tasks and resources that will help facilitate your outcomes assessment planning.

→ Work on the questions with your colleagues.

→ If you see this icon, access the resources for further information and examples by clicking on the link provided.
Assessment cycle

1. Situating
   - Consider organizational readiness for assessment.
   - Determine outcomes assessment needs.

2. Planning
   - Prioritize assessment purposes, uses, SLOs, & questions.
   - Determine indicators for each question.
   - Choose data collection methods & instruments.
   - Decide on analysis & interpretation procedures.
   - Clarify timeline & responsibilities.

3. Implementing
   - Gather data as planned.
   - Monitor data collection process.
   - Analyze data and interpret findings.

4. Utilizing
   - Report findings and discuss ways to make program changes.
   - Prioritize a plan of action.
   - Implement immediate and long-term action plans.
   - Initiate the next cycle of assessment.

   4 phases
Because utilization-focused assessment is intended to respond to the needs of the program stakeholders, we begin by analyzing local assessment needs and readiness to take on assessment work.
Is your department/program ready to proceed with SLO assessment? Many evaluators have found that organizations need to be prepared to use assessment findings to improve and strengthen their program (Preskill & Torres, 1999). Below are some organizational/program characteristics that enable programs to launch, implement, and utilize assessment in a beneficial manner. If you don’t have these in place, consider how you can nurture or acquire them while you engage in assessment.

- Leadership commitment to SLO assessment (e.g., advocacy, support)
- Dedicated committee and/or a coordinator for SLO assessment
- Culture of and experience with faculty collaboration
- Faculty’s openness to change and improvement
- Resources (tools and fiscal support) for conducting SLO assessment
- Shared language and understanding of SLO assessment
- Expertise in assessment (If none, support for faculty development on assessment)
- Reward, recognition, and acknowledged merit for doing assessment work
- Consensus-based outcomes statements and curriculum map
Situational analysis: Assessment needs

In the face of accountability, multiple needs, motivation, and demands for assessment surround a FL program. It is important to consider how your program might negotiate potentially conflicting impetuses and needs, so that you can effectively manage internal and external demands and make the most out of assessment efforts.

• Are there any internally-motivated reasons to conduct outcomes assessment?

• Are there any external requirements or pressure for conducting outcomes assessment (e.g., accreditation-driven outcomes assessment)?

• How would you balance internal and external needs and demands for outcomes assessment?
What SLO assessment needs and demands does your program face? **Who** needs to know about student learning? **What** information needs do they have? **What intended uses** do they have in mind?

<table>
<thead>
<tr>
<th>Who?</th>
<th>What?</th>
<th>Intended uses?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Students’ future employers</td>
<td>Graduating students’ credentials and skills</td>
<td>To recruit appropriate and qualified graduates</td>
</tr>
</tbody>
</table>

Are there any overlaps in the assessment needs and demands of different parties? Are any needs/demands especially important or deserving of attention? If not, prioritize assessment needs and/or (if possible) think of ways to meet multiple assessment needs and requirements.
Once you have identified capacity needs and assessment demands, move onto the next stage: Planning assessment.
Various purposes & intended uses

It is important to consider outcomes assessment purposes with an assessment committee or colleagues in your program before jumping into the methods of assessment. For each purpose, there may be numerous specific uses.

**FORMATIVE**
- Improve, develop, advocate

**SUMMATIVE**
- Hold accountable, demonstrate value, judge

**ILLUMINATIVE**
- Understand learning, understand programs, test theory

**PROCESS**
- Motivate, empower, educate
Various purposes & intended uses

It is important to consider outcomes assessment purposes with an assessment committee or colleagues in your program before jumping into the methods of assessment. For each purpose, there may be numerous specific uses.

**FORMATIVE**
Example use: By identifying which skills, knowledge, and dispositions didn’t meet the expected performance criteria, we can improve and develop course content and tasks to address the outcomes that were poorly achieved.

**SUMMATIVE**
Example use: By using student assessment information, we can demonstrate and justify the value of the program to the university administration.

**ILLUMINATIVE**
Example use: By reviewing progress of students, we can shed light on how students are learning and what works in terms of teaching and learning. The more faculty understand the nature of student progress, the better we can organize pedagogy.

**PROCESS**
Example use: By engaging learners in the assessment process, the process itself is used to unveil and advocate learners’ views on their own learning. Assessment will also be used to enlighten the learners by revealing how much (and what) they have learned in the program.
Key to sustainable outcomes assessment practice is feasibility. When planning for outcomes assessment, you may come up with a long list of possible intended uses and numerous program-level SLOs. The problem is that each intended use and SLO may entail various assessment methodologies, and the scope of trying to do it all may seem daunting. In order not to overburden the faculty and ensure feasible assessment that leads to actions, meaningful prioritization should be embedded at every stage of assessment from design through to utilization (Davis, Sinicrope, & Watanabe, 2009).
In slide 9, you have identified potential intended uses of assessment by stakeholder groups. Among the intended uses articulated, which assessment use should be prioritized right now, and why?

PRIORITIZE INTENDED USE → Among the brainstormed assessment purposes and uses in the previous task, consider what you should prioritize for the first round of assessment. Remember that assessment is a cyclical practice and a small, incremental assessment focus will increase the likelihood of assessment utilization.

PRIOTITIZE OUTCOMES → Not all SLOs need to be assessed at the same time. By focusing attention on high priority needs and uses of particular outcomes, you can make the assessment process feasible, practical, and meaningful for the primary intended users.

Among the program-level or subprogram-level outcomes, which outcome(s) should be prioritized right now, and why?
Assessment Questions

The end product of the intended use and outcomes prioritization process is a series of questions to be answered through assessment. Before deciding how to gather evidence, consider **what** specifically you want to know.

- Assessment questions are guiding questions that provide structure and focus to assessment design and steer data gathering and instrumentation.

- All subsequent steps in the assessment process, particularly assessment methods, must be designed to answer assessment questions.
Developing assessment questions can be one of the most challenging parts of an assessment exercise, so consider carefully how you formulate the questions with your colleagues. Below are some caveats for formulating a good assessment question.

**Caveats for forming assessment questions**
- Questions (Qs) need to be empirical and specific.
- Qs should be realistically answerable.
- Answers to the Qs shouldn’t be known already.
- Users should really want to answer the Qs.
- Users should be able to specify the relevance of an answer for specific future actions.
Assessment Questions: Examples

1. Process questions

- Were intended, sufficient, and high quality learning opportunities provided to the target students?
- To what extent are students satisfied with the program?
- What are the instructors doing that is (not) working well?

2. Outcomes questions

- Is student progress at a certain curricular juncture as expected?
- What do graduating students think they can/cannot do by the end of the program?
- Are the outcomes valuable to students and others?

As you can see from these examples, assessment questions can be about the process of language learning (when, how, factors affecting learning) or outcomes (what and how well) of student learning. For outcomes questions, depending what you need to find out, they may entail perception or performance data.
Before moving on to assessment methods, let’s step back and think about the information needed to answer the assessment questions. The needed information is called an “indicator.” Indicators are the evidence or phenomena that will be useful in answering the assessment questions. An indicator describes a specific way of observing the phenomena of interest.

For example, in order to respond to the question, “To what extent are students satisfied with the program,” what kind of indicators can you think of?

- student perceptions
- retention rate (i.e., ratio of students continuing to the next semester)
- persistence rate (i.e., ratio of students continuing as majors)

→ Some of these indicators are not direct or can be affected by many other factors, and may not be valued by faculty. It is important that the indicator chosen is valued and perceived as interpretable and accurate by the intended users.
**Assessment questions to indicators**

Were intended, sufficient, and high quality learning opportunities provided to the students?

To what extent were students engaged in the learning activities?

To what extent are students achieving the following: Using foreign language skills to locate and use resources (e.g., foreign language texts) in various disciplines.

- e.g., teaching and learning activity logs, teaching materials, teaching performance, student perception of teaching quality and learning opportunities
- e.g., participation rates, student and teacher perceptions of student engagement
- e.g., written and/or spoken research reports that are based on target culture/language resources
Prioritize again!
--Focus, focus, focus--

Assessment questions
You can see from the examples in the previous slide (slide #18) that one assessment question may require multiple indicators. Consider prioritizing assessment questions by…

• critical issues or concerns
• immediate decisions that need to be made
• interest and meaningfulness

Indicators
Also, one indicator may require multiple methods of data collection. Therefore, consider which indicator(s) will most likely provide intended users with an accurate depiction of learning process or outcomes, and prioritize them based on the likelihood of data utility.
Prioritization strategies

When you anticipate contentious discussions or marginalization of voices (e.g., at a meeting), we recommend gathering opinions anonymously prior to meeting via a survey or using a structured meeting facilitation approach. Below are two such prioritization strategies that are commonly used in evaluation studies. Click the link to learn more about the strategies.

- Delphi
- Nominal group technique
Let’s take a look at a real-world outcomes assessment case from Duke University’s FL requirement program (Walther, 2009). Over the next two slides there is an overview of the program context, primary intended users, intended uses, and assessment questions. First, consider the description of the assessment setting below.

1. **Program context**
   - Duke University, general education curriculum for the School of Arts and Sciences.
   - A FL requirement was instituted to develop cross-cultural fluency, understanding, and competencies needed for successful interaction. 20 languages are offered across five departments.

2. **Program focus**
   - Language requirement program across languages

3. **Primary intended users**
   - Provost of undergraduate education (initiator), four language directors (support from the Office of Assessment)
4. **SLOs under focus: Inter-cultural SLOs**
   - **SLO 1**: Gain an understanding of the nature of culture in as far as it is embodied in language.
   - **SLO 2**: Bring a cultural perspective to bear to enhance understanding of issues of similarity and difference.

5. **Assessment purposes**
   - **Formative**: programmatic and curricular improvements
   - **Demonstrate**: program values

6. **Intended assessment uses**
   - **To better understand** the extent to which FL programs prepare students to meet the goals and objectives articulated by the curriculum;
   - **To be able to demonstrate** the merit and value of the foreign language requirement in meeting its stated goals;
   - **To enhance the profile** of foreign language programs within the university at large, as integral to the humanities and to a liberal arts education.
Now, based on the assessment questions below, what indicators would you choose to examine?

**Assessment Question 1:**
What cultural knowledge, understandings, and perspectives do students gain at each respective level of the curriculum?

**Assessment Question 2:**
To what extent do factors such as study abroad, heritage, and previous language learning experiences play a role in students’ cultural learning outcomes?

**Assessment Question 3:**
How do students themselves perceive the impact of the foreign language requirement, in terms of their gains in cultural knowledge and understanding?

→ Go to the next page for Duke’s answers.
Based on the assessment questions, what are the variety of indicators that need to be elicited?

**Q1: Indicator(s)**
- Student attitudes and perceptions on culture.
- Student knowledge of the target culture, society, history, and cultural products.
- Students’ perception of their competency with respect to cultural knowledge and understanding.

**Q2: Indicator(s)**
- Student background info: Time spent living, working, or studying abroad.
- Identity as a heritage or native speaker of the language.
- Language learning experience outside of Duke and prior to taking Duke FL courses.

**Q3: Indicator(s)**
- Student perception, dispositions, and future plans with regard to FL learning (e.g., increase in interest in language and cultural activities; seeking out interaction)

**Future info gathering:**
- Alumni perception on the value of FL study and impact in their career.
- Alumn’s use of FL at their workplace.

Duke FL requirement program
Once you have identified indicators you would like to gather, it’s time to consider how you are going to gather data related to them. There are many ways to gather data on student learning. One major challenge for FL educators is to know which assessment method, among the variety of available alternatives, to choose in telling a particular story about their students.

Consider: What are some of the data collection methods that can shed light on student learning and learning experiences in FL education?

→ Go to the next slide.
Variety of methods

- Direct
  - qualifying exams
  - thesis, dissertation, publications
  - public presentations
  - standardized tests (e.g., SOPI, Dialang)
  - portfolios
  - course-embedded assignments, tests, performances
  - capstone projects

- Indirect
  - exit interview (perception or performance)
  - self reports (logs, journals, self assessment)
  - case studies
  - focus groups
  - surveys
  - observations

→ Provide different pictures
→ Suited for different information needs
→ Complement each other
Methods decisions are never purely technical and are always constrained by…

- limited resources and time for instrumentation and data analysis
- access to informants and data sources
- competing and conflicting priorities
- pragmatic considerations
- disciplinary biases and measurement traditions

In your assessment working group, consider strengths and weaknesses of potential methods and identify methods that will provide useful information about your target indicators. Negotiate methodological choices and determining credible evidence. To help facilitate the methodological decisions, download and use the methods handout. Don’t forget to take advantage of the instruments you already have in place!
Let’s take a look again at the assessment example from the Duke University FL requirement program. To elicit information on the indicators listed below, what kind of data collection tools and methodologies would you choose?

### Q1: Indicator(s)
- Student **attitudes** and perceptions on culture.
- Student **knowledge** of the target culture, society, history, and cultural products.
- Students’ **perception** of their competency with respect to cultural knowledge and understanding.

### Q2: Indicator(s)
Student background info:
- **Time spent** living, working, or studying abroad.
- **Identity** as a heritage or native speaker of the language.
- **Language learning experience** outside of Duke and prior to taking Duke FL courses.

### Q3: Indicator(s)
- Student **perception, dispositions, and future plans** with regard to FL learning (e.g., increase in interest in language and cultural activities; seeking out interaction)
**Method(s)**
(a) Intercultural Development Inventory (A 50-item 5pt-scale survey based on Bennett’s model)
(b) Global Perspectives Inventory (A 46-statement survey, cognitive, intrapersonal, and interpersonal dimension of global perspectives)
(c) Student survey

**Informant & sampling?**
Students who completed FL requirement & students who have not completed the requirement (1/2: IDI, 1/2: GPI)

**Timing?**
At the end of the semester

---

**Method(s)**
(a) Student questionnaire
(b) Information gathering from admissions, study abroad office, and registrar on student background (Office of Assessment collected this info.)

**Informant & sampling?**
(a) Pilot post-course survey: students in the three courses of 4 languages, completing the requirement (inter I, II, & adv-level) in fall 2007
(b) On-going (demographics)

**Timing?**
(a) Post-course (survey)

---

**Method(s)**
(a) Student questionnaire
(b) Information gathering from admissions, study abroad office, and registrar on student background (Office of Assessment collected this info.)

**Informant & sampling?**
(a) Pilot post-course survey with students in the three courses of 4 languages, completing the requirement (inter I, II, and adv-level), during fall 2007

**Timing?**
Revised survey: pre-post survey spring & fall 2008
Duke FL requirement program

Are there any additional methods you would consider?

### Q1: Indicator(s)
- Student **attitudes and perceptions** on culture.
- Student **knowledge** of the target culture, society, history, and cultural products.
- Students’ **perception** of their competency with respect to cultural knowledge and understanding.

### Q2: Indicator(s)
Student **background** info:
- Time spent living, working, or studying abroad.
- Identity as a heritage or native speaker of the language.
- Language learning experience outside of Duke and prior to taking Duke FL courses.

### Q3: Indicator(s)
- Student **perception, dispositions, and future plans** with regard to FL learning (e.g., increase in interest in language and cultural activities; seeking out interaction)

### Method(s)
- Course embedded students’ performance materials via student portfolio
- Self-assessments of cultural and intercultural learning.

### Method(s)
None

### Method(s)
- Focus groups with current students on their experience abroad
- Focus groups with alumni on FL use in their career.
Guidelines for methodological choices

1. Review available data collection methods and consider strengths and weaknesses. Identify potential methods that will provide useful information about your target indicators.

2. Are existing data sources sufficient for meeting some of the information needs? How might they be revised to provide useful information for your specific assessment uses? Take advantage of what you already have!

3. Consider technical, time, and resource demands imposed by different methods, then select methods that meet your needs.

4. Provide clear and reasonable justifications for choosing each methodology. Discuss the merits of your methodological selections with an assessment specialist and with your colleagues.
Guidelines for methodological choices

5. Begin planning for the development of data collection tools and procedures. What resources do you need and what is the timeline? Who are the informants?

6. Is there technical expertise in your program sufficient for meeting the demands of distinct data elicitation methods? What steps might be taken to compensate for the technical demands?

7. Do you need to pilot-test the instrument to make sure that the instrument is reliable and valid?
Methodological choices

Download the worksheet and plan your assessment data collection methods with your colleagues (p. 1).

**Foreign language assessment directory:** The Center for Applied Linguistics (CAL) has a directory of foreign language tests (nearly 200 tests in over 90 languages).

“Understanding Assessment: A guide for foreign language educators” This is an online tutorial on language testing by CAL. The module covers practicality, reliability, validity, and impact of testing.

University of Texas at Austin, Instructional Assessment Resources has a comprehensive website on methods selection for program evaluation. [http://www.utexas.edu/academic/ctl/assessment/iar/programs/plan/method/](http://www.utexas.edu/academic/ctl/assessment/iar/programs/plan/method/)

Visit our Foreign Language Program Evaluation website. We have a collection of online guides and resources on program evaluation and outcomes assessment. [http://www.nflrc.hawaii.edu/evaluation/resources.htm](http://www.nflrc.hawaii.edu/evaluation/resources.htm)
Methods and Instruments: Validity

Recall that a key aspect of assessments is the usefulness of information generated through data collection. Unless assessment findings lead to meaningful interpretations of student learning, assessment use for decisions and actions will not happen. Therefore, when creating or adapting assessment methods and instruments, ask the following questions:

• What learning characteristics is the instrument (questions, prompts, tasks, protocols) trying to capture? Do the characteristics captured in the instrument align with the outcomes? Do they provide necessary information?
• What inferences can we make about learning from student responses, performances, or observations? Do we trust the instruments?
• Does the instrument/method align with what and how students have learned?
In order to achieve a good alignment between what gets assessed and what we can interpret about learning outcomes, we need to articulate specific performance characteristics and/or expectations (i.e., learner profiles). For example, if we were to create an instrument to capture the following SLO, how might we define our expectations for learner performances? What would a learner profile look like for this outcome at this curricular level?

*From* [Georgetown University German Department](https://www.georgetown.edu/german/) *(Level I)*
By the end of the course students will be able to understand and produce a variety of text types, from personal and interactional to routine public, supported by reading and listening development.
We might use sources like the following to come up with a performance profile: (a) writing samples that showcase the ability of the target outcome; (b) prompts and tasks that are in use to examine what learners can produce; (c) language proficiency guidelines and standards; (d) teachers’ descriptions of what learners can do.

Here is an example learner performance profile for curricular Level I at the Georgetown University German Department: [click here]
When examining the level of student achievement, satisfaction, or confidence (outcomes questions), judgment criteria, also known as “benchmarks”, need to be agreed upon by the intended users. The criteria provide meaning to a rating or a score—how much is good enough?—and lend coherence and clarity to expected learning profiles. Determining judgment criteria (i.e., standard setting) is value-laden and political; thus, careful consideration is needed in deciding “who provides the judgments for setting the performance standards and the context in which those judgments are provided and interpreted” (Hambleton, 2001, p. 93).
1. Internal or external standards?
   In order to choose an appropriate benchmark, we can (a) refer to standards set by disciplinary associations, such as language-specific teaching associations or the national FL standards; (b) involve stakeholders (faculty, students, etc.) in standards-setting processes; and/or (c) make use of existing or newly gathered student work samples to determine what is an adequate performance.

2. Rubric development
   Typically, standards take the form of a rubric that describes different levels of performance. To create a rubric, you need to decide whether it should be...
   - Generic or task-specific
     *Is the rubric going to be applied for a specific task or across tasks?*
   - Analytic, holistic (or checklist)
     *Will the focus be on different characteristics of a performance or the overall quality of a performance?*
### Components of a rubric

- **Scale:** Typically, 3-6 levels. Consider the appropriate number of levels in alignment with your program’s understanding of students’ developmental patterns (e.g., by curricular semester/year).

- **Performance attributes:** Determine the features of a good task performance, or the types of competence necessary to perform the task.

- **Performance descriptors:** What is the quality of performance for each attribute? Are there developmental expectations for each performance attribute? Define and describe.

#### Descriptions of the quality of performance

<table>
<thead>
<tr>
<th>Performance attribute 1</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance attribute 2</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performance attribute 3</th>
<th>Level 1</th>
<th>Level 2</th>
<th>Level 3</th>
<th>Level 4</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Example:** content knowledge, organization, reflectivity, grammar, mechanics, etc.)

**Scale examples:**
- advanced, intermediate high, intermediate, novice
- exceeds expectations, meets expectations, needs improvement
- exemplary, proficient, marginal, unacceptable
- distinguished, proficient, intermediate, novice
- accomplished, average, developing, beginning
- mastery, partial mastery, progressing, emerging
## Components of a rubric

An example holistic rubric for spoken presentations from the Center for Advanced Research on Language Acquisition, University of Minnesota:

<table>
<thead>
<tr>
<th>Level 3</th>
<th>Student presents all relevant information clearly with a wide variety of sentence structures, grammatical forms, and specific vocabulary. Presentation contributes to the understanding of the intended audience. There may be minor errors that do not impact comprehensibility.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level 2</td>
<td>Student presents most relevant information clearly with a variety of sentence structures, grammatical forms, and relevant vocabulary. Presentation may be less appropriate for intended audience. Several errors may impact comprehensibility, but misunderstandings are clarified in the target language.</td>
</tr>
<tr>
<td>Level 1</td>
<td>Student’s presentation has major gaps in information. Speech is simple and lacks variety. Vocabulary is limited. Presentation does not meet the needs of the intended audience, and more frequent errors lead to confusion.</td>
</tr>
</tbody>
</table>
In order to provide consistent and reliable assessments, those who will be looking at student work or performances need to be familiar with the rubric and able to interpret and apply it in the same way. In order to calibrate ratings among raters, an orientation session can be important. Below are steps involved in rater training and calibration.

**Step 1:** Explain how to use the rubric. Familiarize faculty with the categories and levels. For each mastery level, provide one sample with annotations of the features found in student work that capture the rating criteria.

**Step 2:** Provide two samples of student performance/work that represent different levels of mastery (don’t disclose the ratings). Have faculty rate them independently.

**Step 3:** Gather their ratings to show agreement/disagreement.

**Step 4:** Discuss scoring inconsistencies and justifications. Clarify the rubric, and try steps 2-3 again as needed.

**Step 5:** Once consensus is achieved on the ratings and when faculty feel comfortable with the rubric, proceed with individual ratings of student work/performance (e.g., two raters for each sample). Calculate rater reliability to make sure final ratings are consistent.
After collecting information, the data must be organized in a way that is analyzable and interpretable to the intended users of assessment. The goal of data analysis is to provide order and structure for making sense of the information. When planning for data analysis, start by reviewing the assessment questions. The key consideration is to whose eyes the analysis process and the results have to be credible and trustworthy.
Data analysis procedures: Guidelines

- Who should be involved in data analysis? Are strategies needed here to avoid suspicion of data manipulation?
- How can data be extracted and organized from assessment scores or other results?
- What data analysis techniques should be used? Think of the most direct ways to answer the assessment questions.
- How can reliability/trustworthiness of analyses be maximized? For example, for student writing data, reliability can be assured by using multiple raters (e.g., two raters for each dataset).
Based on the results of data analysis, interpretations and value judgments have to be made in order to respond to assessment questions and lead to uses. Since personal perspectives often affect data interpretation, careful consideration of this stage will be necessary. Here are some guiding questions:

- Whose perspective counts? Who should be involved in data interpretation? Why?
- How will interpretation be checked? Is triangulation of sources and perspectives necessary?
- To what extent should findings lead to implications for program change?
- Who gets to draw implications, and who gets to make recommendations?
- Are the interpretations and recommendations evident from the data and clearly articulated in understandable language for the intended users?
Recommended resources for data analysis & interpretation

- University of Texas at Austin: Step-by-step guide to analyzing and interpreting data for various methods (document analysis, focus group, interview, observation, survey).
  Visit → http://www.utexas.edu/academic/diia/assessment/iar/programs/?step=report

- University of Wisconsin Cooperative Extension: Detailed reader-friendly guide to qualitative and quantitative data analysis methods.


Download the worksheet and plan your data analysis and interpretation with your colleagues (p. 2).
Now that you have a work plan, the next stage is putting all the planned assessment activities into action. By monitoring implementation, factors that may affect assessment findings can be identified and controlled along the way.
Often, assessment may not go as initially planned, due to contextual constraints or overly ambitious plans. However, changes in initial planning can affect intended use of the findings; thus, we need to make sure that a responsible person or group monitors the entire assessment implementation process.

Monitoring assessment implementation requires management skills. The person monitoring data collection needs to (a) engage primary intended users in the process and provide updates; (b) ensure credibility and integrity of assessment practice; (c) deal with unexpected challenges; and (d) seek opportunities for capacity building when necessary.
Implementing

Monitoring: Data gathering, analysis, & interpretation

Guiding questions:

• Create opportunities to reflect on the data collection processes to examine whether there are any potential threats to reliability and validity of data.
  o Were the instruments implemented as planned?
  o Did anything go wrong or different than expected?
  o Do you have sufficient data to come to conclusions?
  o Do you think that the data is representative of the students you have?

• Manage data carefully, so that reliability and credibility of data are assured.
Implementing

Monitoring: Data gathering, analysis, & interpretation

- Distinguish interpretations, judgments, and recommendations.
  - **Interpretation** “involves determining the significance of and explanations for the findings”
  - **Judgment** “brings values to bear to determine merit, worth, and significance, including the extent to which the results are positive or negative.”
  - **Recommendations** “involve determining the action implications of the findings.” (Patton, 2012, p. 337)
Finally, we are at the stage of closing the loop of assessment: the utilization stage. Note that dissemination and reporting do not entail utilization, although they are required if use is going to happen. Use should go beyond reporting to taking action.
Reporting assessment findings and recommendations is also guided by the principle of intended use by intended users.

**Caveats**

- Be intentional about reporting (know the purpose/use of a report)!
- Be user- and audience-focused!
- Don’t surprise the users.
- Think positive about negatives
- Distinguish dissemination from use
• Who are the potential audiences?

• Who contributes to reporting?
  Collaborative?
  At what point do you involve them?

• What do assessment users expect from reporting?
Reporting: Format?

- Meeting
- Workshop
- Presentation (online, face-to-face, video)
- Poster
- Website report
- Brochure/newsletter
- Press release
- Monograph/article
- Executive summary
- Full written report

- More interactive
- Less interactive

☑ Access?
☑ Engagement?
☑ Use?
Reporting: Different audience, different format?

- Meeting
- Workshop
- Presentation (online, face-to-face, video)
- Poster
- Website report
- Brochure/newsletter
- Press release
- Monograph/article
- Executive summary
- Full report

Program-internal faculty?
Students in the program?
University admin?
Potential students/parents?
Peer institutions/program?
Alumni?
Reporting: Guiding questions

- What is the **purpose** of reporting? (multiple purposes?)
- Who is the **audience** of reporting? (multiple audiences?)
- What steps do you want to take for interpretation and reporting?
  - Who should be involved in the process?
  - Collaborative or non-collaborative?
  - More interactive or less interactive?
  - How much **time** do you think the process will take?
- What **additional resources** would help in addition to the report?
- What kind of **results** were anticipated?
- What are some of the **contextual constraints** in taking further action, and what strategies can you adopt to move forward?
Resources on effective reporting

(1) A useful book on reporting evaluation results and findings:

(2) Simple guide on good data reporting practices:

(3) Many links about doing good presentation/reporting of results:
   [http://gsociology.icaap.org/methods/presenting.htm](http://gsociology.icaap.org/methods/presenting.htm)

(4) Analysis and reporting of survey data:
   [http://www.rdg.ac.uk/ssc/publications/guides/topasd.html](http://www.rdg.ac.uk/ssc/publications/guides/topasd.html)
   [http://www.rdg.ac.uk/ssc/publications/guides/toptgs.html](http://www.rdg.ac.uk/ssc/publications/guides/toptgs.html)

(5) A checklist on writing an evaluation report
   Kiernan, Nancy Ellen (2009). An evaluation report. TIPSHEET #91, University Park, PA: Penn State Cooperative Extension. Available at:
   [http://www.extension.psu.edu/evaluation/pdf/TS91.pdf](http://www.extension.psu.edu/evaluation/pdf/TS91.pdf)
Do not assume that what is reported and disseminated will automatically be used by the intended users. Are the intended users capable of utilizing the findings? Utilizing the findings may require training and careful planning.

- Is an action plan needed?
- Is action already in place?
- What about the next evaluation cycle?
Action planning

Steps:
1. Review original assessment plan, and findings thus far, and determine what needs to happen next. What actions would you take based on the results?

2. Select a reasonable target date and plan backwards, considering likely time necessary for accomplishing discrete objectives.

3. Determine a feasible and concrete timeline of events. What additional resources are needed to take desired actions?

4. What are some of the contextual constraints in taking an action, and what strategies can you adopt to move forward?

5. Consider who is going to be involved and at what point in time for making program changes. Clarify responsibilities: who needs to do what when?
Factors that influence taking action

Ownership of findings
Feasibility of actions
Accessibility of findings
New capacities through assessment
Increased control over evaluation
Sufficient guidance for future action

Anticipate and try to enhance these factors.
By going through one assessment cycle and closing the assessment loop, you may find that additional assessment questions emerge. Keep track of new questions and address them as you plan for the next assessment cycle.

Guiding Questions:
• Based on the assessment findings, did you find any changes necessary in the student learning outcomes themselves? Are they reasonable, feasible, appropriate?
• Do you need to collect more data in order to answer the assessment questions?
• Did additional questions come up during the assessment process?
• Which SLO or assessment use needs attention next?
Plan for utilization

Complete the assessment planning worksheet and plan for assessment reporting and action-taking (p. 2).
Conclusion

Section IV covered the entire cycle of doing outcomes assessment from situating to utilizing. Assessment is a learning process and should remain so, even for those who have been engaged in assessment for some time. Going through one cycle of collective inquiry should enable programs to determine how to fit assessment practices into organizational routines. The last module in this series, Module V, suggests points for reflection on the entire outcomes assessment process in order to improve the contribution of assessment itself to our FL educational endeavors.