2011 Middle Eastern Languages Western Consortium  
Language Program Evaluation Workshop 
Roundtable Discussion #2, Saturday, July 30, 2011, 4:45 – 6:30  
‘Assessing’ and otherwise gathering data on diverse program outcomes: Moving beyond ‘how do we measure?’

Participants: John Norris, Martha Shulte-Nafeh, Esther Raizen, Ahmet Okal

Summary: Evaluation calls upon empirical data as a primary basis for informing decisions and taking actions in language programs. Yet there are numerous possible methodologies for gathering data, from assessments of student learning, to observations of how well programs are delivered, to perceptions of satisfaction and impact. Indeed, many of the outcomes associated with language programs may defy easy ‘measurement’. In this roundtable discussion, participants will provide insights into useful methods for collecting meaningful data on the distinct kinds of outcomes (learning and otherwise) that MELPs seek to achieve.

1. Taking our programs seriously: From outcomes to indicators

- Generally speaking, outcomes embody the valued changes or benefits that accrue—to learners, groups, institutions, scholarly disciplines, even to society in general—as a result of the activities, services, courses, and other efforts we make in language programs.
- Typical outcomes in MELPs may focus on what students learn or develop, such as targeted levels of language proficiency, the mastery of particular bodies of literary or other cultural knowledge, ability to study or do research or work in a Middle Eastern linguistic-cultural context, and others.
- Most language programs also promote other kinds of potentially less obvious outcomes. For students, these may include changes in intercultural competence, aesthetic sensitivity, critical self-awareness, and other ‘esoteric’, humanistic kinds of learning. For language/culture programs, these may range from the concrete and specific, such as the publication of a certain number of materials or the successful hosting of cultural events, to the abstract, such as raising awareness about ME languages/cultures among particular constituencies or achieving discernible changes in public opinion.

Notes:

- Data on outcomes provide one of the primary means for understanding program effectiveness and impact. However, when it comes time to collect outcomes data, a common reaction from is that “Many of the outcomes we target cannot be measured”.
- An alternative starting point may be to ask “What does the outcome look like?” and “How might we know that the outcome has been achieved?” Thus, instead of beginning with the assumption that all outcomes must be measured, it may be more productive to consider the variety of possible types of evidence that can show us something about the actual outcomes that our programs target. These various types of evidence can also be thought of as “indicators”, that is, behaviors, reactions, opinions, or other observable phenomena that indicate something meaningful about an outcome.

2. What are some of the possibilities for assessing and otherwise collecting data on challenging outcomes?

- Direct methods generally involve the observation of learners’ or participants’ actual behaviors, either during program activities or as a result of them. Direct methods typically answer the overarchig question of “What can/do participants do as a result of the program?” For example, in a study abroad program that seeks to develop participants’ willingness and ability to engage in rational discussions on sensitive cross-cultural topics, we might: (a) have participants perform simulated discussions at various points (beginning, middle, end) in class, with ratings by peers and an instructor on a jointly constructed “cross-cultural discussion” rubric; (b) have participants keep a log of such discussions actually undertaken outside of class over the course of a stay abroad; and (c) have pairs of participants alternately engage in ‘person-on-the-street’ discussions while a peer observes/rates how well the discussion takes place.
- Indirect methods involve the elicitation of perspectives, opinions, or other self-report data in order to find out about outcomes that may not be easily, efficiently, or thoroughly observed in behaviors. Indirect methods can answer the “what” question above, but they are also used to address “Why do participants behave in a certain way as a result of the program?” For example, in the above scenario, we might follow up with participants after the conclusion of the study abroad program by administering an anonymous survey that asks them to: (a) provide an honest assessment of their perceived development in cross-cultural discussions, both in terms of willingness and ability to engage; (b) comment on key cross-cultural learning moments during the stay abroad; and (c) critique the effectiveness of in-class preparation for achieving this outcome. Patterns of responses from participants might help to indicate how well the program experiences, both in class and outside of class, prepared students to engage in sensitive cross-cultural discussions.
• The key with gathering data on challenging outcomes is to think first about what the outcome might look like, how it might be observed, or how critical information about it might be elicited in insightful ways (i.e., in ways that are relevant to the values and interests of our programs, participants, and scholarly disciplines).
• Being honest to the actual outcomes we want to target often means that we must be creative in designing data collection methods that will tell us something useful about them.

3. How do we make use of findings about outcomes? A few ideas.

→ Ultimately, how we choose to assess student learning, or to otherwise gather data on MELP outcomes, will depend a lot on what we intend to do with the findings. Having a good idea about the intended uses for outcomes data will help to narrow down the appropriate methods.

→ Outcomes assessments are increasingly called upon in language programs to help demonstrate the worth or value of participation, and ultimately the tangible impacts of the program. Thus, in MELPs we may need to engage in public relations, recruitment of participants, even defending of our programs from budget cuts. The types of data on outcomes that will prove effective for these more public uses may differ considerably from the data we would collect for internal purposes. Here, the key may be to think in terms of what kinds of information various actual audiences will be able to understand and respond to. For example, while learning portfolios of students’ development in writing and speaking abilities, content mastery, and critical thinking over four years of language study may give us rich data for inquiring into actual achievement of outcomes throughout a series of courses, it may be pretty difficult to convey to a non-teaching public exactly what a portfolio demonstrates about the ultimate worth or value of the program. Alternative indicators, such as job or graduate school placement, case studies of successful alumni, standardized proficiency test scores, employer assessments of graduates, and so on, may prove more effective at communicating program worth and value on terms that are understandable to a variety of public audiences.

→ Certainly, a major intended use of outcomes data is to identify aspects of our programs in need of improvement. Good evidence about which outcomes are being achieved and to what extent, as well as which are not, can point us in the direction of curriculum, course, materials, outreach, and other areas in need of development work. Key to formative, improvement-oriented information is the collection of data in a way that can link program activities (what we do) to program outcomes (what happens as a result). For example, simply knowing how many individuals participate in a series of ME cultural events sponsored by a language center may tell us something useful about potential impact on the community, but it probably won’t tell us much more.

We may also want to find out exactly how participants thought they benefitted from the events, which aspects of the events helped them to benefit, and which aspects did not. Accordingly, in addition to participation tallies, we may want to administer focus groups with a randomly selected set of participants in order to link their participation to aspects of the program in need of improvement, as well as those that worked well and should be maintained.

→ Articulating and persistent monitoring of outcomes may also serve the important role of needs assessment for ME programs. In the first instance, by forcing ourselves to achieve consensus regarding exactly what set of outcomes represents the valued benefits or impacts offered by our programs, we can begin to come to terms with the societal and/or individual participant needs to which we think our efforts are responding. In an era of shrinking budgets and increased public scrutiny, the ability to state that an array of outcomes (including humanistic outcomes) actually meets the needs of specific segments of society is perhaps not a bad thing. Beyond stating our program’s outcomes, the on-going collection of data regarding them can serve as a perpetual check on both the effectiveness and the relevance of what we are doing. Just as learner or participant populations change, public opinion shifts, and funding waxes and wanes, so too might the needs—to which our programs should be responding—evolve. For example, a simple alumni and employer survey collected over a period of several years might be used to monitor the extent to which the key outcomes we have articulated, and which we seem to be accomplishing, are in fact perceived by graduates to be the right ones to target.

4. Questions for the panelists:

• What are some of the more challenging outcomes to assess in MELPs?
• What assessment (or related) methodologies seem to work best in providing useful information about diverse MELP outcomes? What one or two examples would you like to showcase?
• What are some of the difficulties in analyzing and making sense of assessment (or related) data for diverse outcomes—what else do we need to keep in mind beyond collecting data (e.g., rating scales and procedures)?
• How do assessment (or related) findings about program outcomes get used? What happens as a result of assessment?

Notes: ____________________________________________________________
__________________________________________________________
__________________________________________________________
__________________________________________________________