ED41D-0700: Lessons Learned from IRIS EPO Program Evaluations

Relative benefits of internal and external assessments

JOHN TABER & MICHAEL HUBENTHAL
IRIS Consortium (taber@iris.edu Michael.Hubenthal@iris.edu

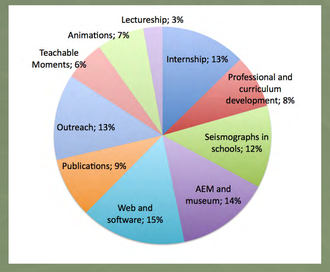
Overview

Evaluating the overall impact of EPO programs that include activities ranging from formal education through broad public outreach, is a complex issue. The impact of education activities targeted at narrowly defined audiences is generally easier to quantify than the national impact of outreach activities conducted by a relatively small program. For educational activities, our approach has been to leverage the best-practices identified through research and to continuously assess the individual elements internally with the intention of making improvements based on the data generated and the existing research. Ongoing review by an EPO advisory committee provides regular oversight of program impact, while targeted but less frequent external formative and summative evaluations also help improve the program and guide decision making.

EPO Program Goals

Clear goals and carefully defined content are key aspects of the program. The EPO program goals are:

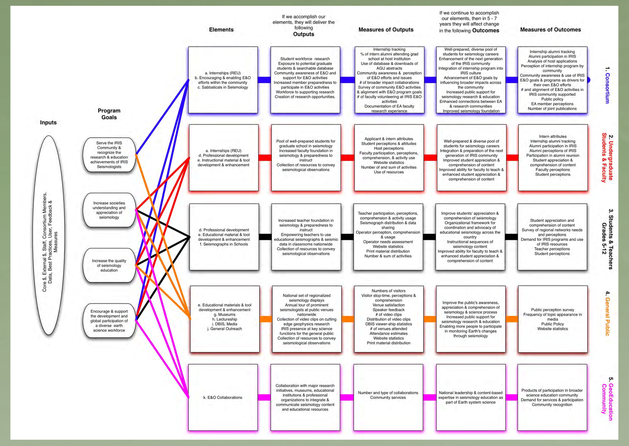
- *Improve Seismology Education:* Increase the quantity and enhance the quality of seismology education
- *Expand Earth Science awareness:* Expand opportunities for the public to under stand and appreciate seismology
- *Support IRIS Consortium members:* Provide education and outreach products and services for members of the IRIS community.
- *Enhance IRIS Visibility:* Increase the visibility and recognition of IRIS through effective branding and communication of IRIS EPO products and services
- *Expand the Earth Sciences workforce:* Support the development of a larger and more diverse Earth science workforce
- *Strengthen the EPO Program:* Seek collaborations and funding to sustain and grow the E&O program.



We have chosen to develop a portfolio of activities that range from short interactions (minutes) with millions of visitors, to an entire summer of research for 15 students. Within this range of activities we work continuously to balance impact and quality versus cost.

The IRIS EPO program has worked to maintain a balance between the education and outreach ends of the program mission. This is a challenging course to maintain because it is often hard to assess the relative worth of an activity that greatly impacts a few people (education) compared to an activity that minimally impacts a great number of people (outreach).

Logic Model



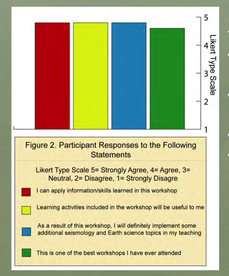
The above logic model was developed as part of an external summative evaluation of the EPO program in 2009. The logic model and associated strategic plan helps drive evaluations and decision-making about potential new activities.

Internal Assessment Strategy

By constructing our elements on the best practices identified by the research community we feel that internal formative evaluation is a valid means to determine if an activity is effective, particularly when the results are compared to similar programs. For example:

- Effective practices of professional development are well documented in the literature.
- This allows us to shape our programs and our evaluations to monitor elements that have been identified as key by the educational research community.
- Such actions allow us to avoid allocating significant resources with the intention of pinning down direct causal relationships between our programs and consumers, when similar interventions (conducted by others) have already shown such relation ships.
- Sharing internal assessment results via white papers and conference posters and talks allows further review and dissemination of the findings.

Formative assessment - Professional Development



Workshop attendees are surveyed at the conclusion of workshops and typically the responses show that teachers were very satisfied with the workshop.

Surveys conducted at the end of the school year following a workshop provide a measure of effectiveness. In one example:

- 74% of respondents report increasing the amount of time spent teaching seismology or related topics in their classroom
- 86% of respondents report using at-least one activity modeled during the workshop upon returning to their classrooms.
- An average of 4.5 modeled activities were used per teacher
- In contrast, only 13% of respondents used activities that were made available to participants in the form of handouts at the workshop but were not modeled during the workshop.

These results are then used to modify subsequent workshops. While we have not directly assessed student learning, positive impacts on student learning of seismology related topics can be inferred from the number of participating educators reporting an increase in the amount of time spent on seismology and the number of activities used.

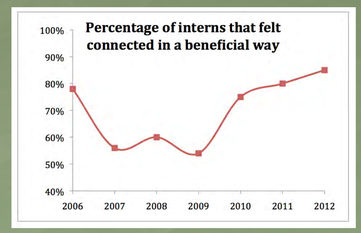
Formative assessment - Research Experiences for Undergraduates

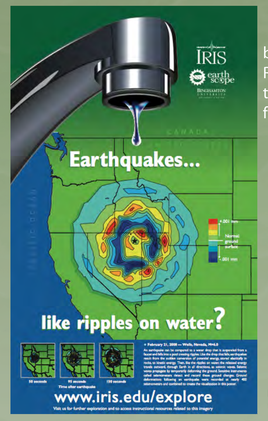
Longitudinal tracking allows long-term evaluation of effectiveness of a particular activity. For example, in our Research Experiences for Undergraduates program:

- 126 students have conducted research with mentors at 47 different institutions over the past 15 years.
- 90% of those who have completed their undergraduate degree have gone on to graduate school or a geo science career.

Annual assessment of the degree to which we have achieved our stated goals for the project allow us to make adjustments to the program to improve outcomes. For example, perception data is used to assess the goal:

- Interns will feel and demonstrate a beneficial connectedness.





This is an important goal for the program because of the distributed nature of our REU, where students spend most of their summers at widely separated institutions. The figure to the right shows the positive effect on student connectedness beginning in 2010 of the switch from an online forum to Facebook as a means of informal communication.

In-Depth Case Study – Educational Posters

This specialized study was driven by the need to better determine the value of creating and distributing large numbers of posters. The study quantified poster usage by teachers in the classroom, determining both the type of posters teachers prefer, and why they like them, which showed that there was a disconnect between posters (theory) and teachers (practice) as conveyers of content. This led to an improved poster design (Hubenthal et al, 2011).

Poster created based on the feedback from teachers highlighting the importance of a central image that the teachers can use to explain a concept.

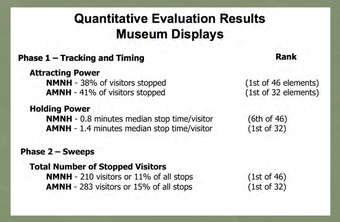
External Evaluations

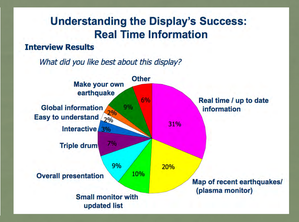
Formative Evaluations

- *Active Earth Monitor content set* - An external evaluator was employed to observe visitors using one museum display content set and to interview visitors about the storyboard for a new content set.
- *Bringing Seismology's Grand Challenges to the Undergraduate Classroom* – independently funded curriculum project with a standard external evaluator and evaluation plan

Summative Evaluation

- IRIS/USGS Museum display - A summative evaluation of the same implementation of a display at two museums allowed us to compare the audience response relative to other displays in each museum, and to determine which elements of the displays were most effective (Smith et al, 2006). This led to the development of the smaller, more flexible Active Earth Monitor.





Summative Evaluation Of Overall Program

While we find internal and external formative evaluation extremely useful in shaping the program and documenting its impact, we also recognize the value of a summative evaluation process. For example, an external summative evaluation of the IRIS EPO program was conducted in 2009, followed by an external panel review, as part of the regular review of IRIS programs.

- The most valuable part of the external evaluation was our preparation, including clarifying the goals of each of the elements of the program, developing a logic model, and assembling and internally evaluating already-collected evaluation data sets.
- The external review panel also provided valuable suggestions for improvement of individual elements.
- However while the external evaluator attempted to define measures of the impact of the overall program, using an evaluation tool called RE-AIM (Glascow et al, 1999), both the review panel and EPO staff found these results unsatisfying because of the large number of assumptions needed.
- This may have been at least partially due to a lack of familiarity with our content area, and we think working more closely with the evaluator throughout the evaluation process could help resolve such issues in the future.

Conclusions

- Internal assessment is a valuable tool for continuously improving the quality of an EPO program
- Aligning activities with best practices allows the inference of student impact without conducting educational research
- Targeted external evaluations help to confirm the value of a program
- Quantifying the ideal balance of activities within an EPO program is a difficult task for an external evaluator
- An ongoing advisory panel provides valuable oversight of an EPO program

References

Glascow, R.E., T. Vogt, and S. Boles, 1999, Evaluating the Public Health Impact of Health Promotion Interventions: The RE-AIM Framework, American Journal of Public Health, Vol 89, No. 9.

Hubenthal, M., T. O'Brien, and J. Taber, 2011, Posters that foster cognition in the classroom: multimedia theory applied to educational posters, Educational Media International, Vol 48, No. 3. 193-207.

Smith, M., J. Taber and M. Hubenthal, 2006, Real-Time seismic displays in museum appeal to the public, EOS, Trans. AGU, 87, no. 8, 85.