

Draft: Do not cite or quote without permission

Expert Panel Review of OERL's Instructional Utility

Dan Zalles, Geneva Haertel
Center for Technology in Learning
SRI International

Paper presented at the annual meeting of the American Educational Research Association, April, 2005, Montréal, Canada. Send correspondence to Geneva Haertel, SRI International, 333 Ravenswood Ave., Menlo Park, CA 94025. This research was supported by contracts from the National Science Foundation (REC-9912172 and NSB-0353574).

Do not cite or quote without permission

Expert Panel Review of OERL's Instructional Utility

Why an Expert Panel?

As part of the site's internal formative evaluation, the OERL team has adapted and integrated several evaluation methodologies to examine patterns of usage, customer satisfaction, and the instructional utility of the site. These methods include surveys of target audiences, usability testing, and an expert panel review. This chapter presents the results from the expert review.

The Online Evaluation Resource Library is a Web site of evaluation resources (<http://oerl.sri.com>). SRI International developed OERL under contract with the National Science Foundation. It is intended to provide evaluators of NSF projects with resources they can use to develop sound evaluation plans, reports, and instruments. The site includes a collection of plans, instruments, and reports (selected from artifacts of NSF-funded projects) and professional development modules on topics such as methodology, sampling, and questionnaire development. To date, the site has been visited by thousands of users from a variety of organizations including institutions of higher education, government, military, and industries.

As a library of evaluation resources, OERL has many potential uses, in addition to its primary role as a support to evaluators of NSF-funded projects. In this expert panel review, we chose to focus on the potential for OERL to be used in graduate evaluation courses as a tool for building the capacity of new and potential evaluators. In addition to being a possible teaching tool, introducing OERL to graduate level evaluation students is a form of dissemination. Before we began our expert panel review, we knew that OERL had been used by some professors, but we wanted to better understand its promise for use in graduate level evaluation courses. An expert panel of five university faculty members responsible for graduate training in evaluation were invited to review the site. The review began in mid-December of 2002 and the phone call conferences took place in January and February of 2003.

The use of expert panel reviews is a widespread practice in federal, state, and local governmental agencies in the United States. Experts are convened to deal with policy issues and program initiatives. Expert panels are used as part of contract and grant-funded projects to provide feedback on work conducted. Panels can generate ideas to move a project forward, extend the thinking behind proposals for further work, or improve the technical features of program initiatives.

The rationale for this approach is that if a seasoned, knowledgeable group of professionals is engaged in discourse on a substantive or technical issue, the best ideas on the issue are more likely to be generated than by other means of deliberation. The approach is cost effective in that panel members meet in one location (or at one time, if the panel is convened by conference call) and, during a relatively short period of time, focus intensively on the issue(s) at hand. Members of the panel can be given background materials, prior to their meeting to ensure the efficiency of the process.

A primary purpose of the expert panel review of OERL was to understand what the panelists would characterize as an ideal Web site for the education of their graduate students in evaluation courses. Their ideal could be compared to the design of OERL to see to what extent they align in terms of purpose and audience. In other words, the expert panel review process would not only provide an opportunity to gather commentary on the Web site, but to carry out a needs assessment as well.

OERL Design Decisions

Before proceeding with the panelists' feedback, it is important to understand the origins of key OERL functions and characteristics. In designing OERL, we had to make decisions regarding scope and structure. From its inception, the site was designed to provide a collection of artifacts from actual NSF-funded projects, contact information about the contributing projects, and tutorial help on evaluation methodology issues. The following design issues were addressed.

1. *Should OERL archive all of the materials submitted to it, or should the materials undergo a review process?* The latter course was chosen, on the assumption that only material representing sound evaluation practice, as delineated in the

Program Evaluation Standards (Joint Committee on Standards for Educational Evaluation, 1994), should be posted on the site.

2. *If the complete evaluation plan or report is not of uniform quality, would it be useful to excerpt portions of the document that could be posted on the site?* It was decided that excerpts were of sufficient value to be posted even if the documents in which they are embedded do not represent uniformly sound practices. We would make the excerpts more useful by categorizing and presenting them by evaluation plan and report component, linking them to one of the Evaluation Standards (Joint Committee on Standards for Educational Evaluation, 1994), and annotating them for greater educational meaning. At the same time, a decision was made not to excerpt instruments because the individual items or subscales would be too fragmented and require much more contextualization and Web site infrastructure than had been planned. Hence, instruments would be either accepted or rejected in their entirety.
3. *Should OERL represent a range of evaluation approaches?* It was determined that a goal-driven, outcomes-based approach should be the primary focus, since many evaluations of NSF-funded projects are designed to indicate whether project goals are attained. However, other evaluation approaches would be included if sound artifacts were offered.
4. *How should the content be organized?* Options included organizing by (1) the subject matter of the contributing projects (e.g., calculus, physics, electrical engineering); (2) project type (e.g., curriculum development, teacher education, faculty development, etc.); or (3) quality criteria within an evaluation artifact (i.e., plan, report or instrument). It was determined that the organizing principle would be a combination of options 2 and 3.

These design decisions were not shared with the expert panelists before their reviews, in order to ensure that the panelists would express, without inhibition, their assessment of the Web site and its potential value to their students.

Expert Panel Method

In determining the methodology for the expert panel review, the goal was to maximize the efficiency of the process for getting the most constructive feedback in the timeliest manner. To this end, the following strategies were implemented.

Procedures Used to Conduct the Panel. The panelists were provided instructions for the review task and timelines for completion. Panelists were asked to spend 2.5 days examining the OERL Web site and answering a series of questions concerning the type of evaluation courses they teach, the characteristics of their students, and the utility of OERL for graduate-level evaluation courses. In answering the questions, they were asked to develop written commentaries that were as thorough as those they would develop if reviewing a scholarly book or article.

Given the size and complexity of the OERL site, the panelists were assigned focus areas for the review. Each panelist reviewed two project types and two professional development modules. A review of a project type entailed reviewing plans, instruments, and reports for that type. The work assigned to each panelist partially overlapped that assigned to another, so that two different panelists provided reviews of resources from the same type of projects and the same Professional Development Modules. This allowed for contrasting views and balance in the feedback.

After each panelist completed his or her written review, the OERL team scheduled three teleconferences. The first teleconference was scheduled individually with each panelist. The purpose of this first conference was to give the OERL team the opportunity to hear the panelists one at a time and thereby identify and, if necessary, clarify the features of the site and the intentions of the review process. Representatives from the OERL team asked questions of each panelist concerning his or her written review, as well as other questions that seemed appropriate at the time. This first conference lasted approximately 45 minutes.

The second teleconference included all expert panelists, representatives from the OERL team, and the NSF program officer. During this two-hour teleconference, each

panelist provided an overview of his or her written commentary for the panel and responded to queries from the group. The focus of discussion was the ways OERL could be incorporated into existing courses at respective institutions. To maximize the efficiency of the exchange, all participants were provided beforehand with transcripts and summaries of panelists' written responses to the review questions.

The third teleconference focused on the panelists' review of the project types and professional development modules that they had been asked to review. General feedback on the Web site was also provided. This teleconference was approximately two hours in length.

Design of the Questions. Appendix 2 contains the instructions to the expert panelists. The panelists were asked to identify the target audience of the site. This question required that information about the original purpose of the site be temporarily withheld from them. In addition, the experts were asked to critique the usefulness of the site for the specific courses they taught. They were also given time to express any additional issues that came to their mind when reviewing the site.

The following six questions were presented to each panelist, who was asked to prepare a three-to-four page written response. Prompts followed each question to encourage panelists to address a number of aspects of the site.

1. What kind of courses on evaluation do you teach?
2. How suitable is the site as a resource for your courses in evaluation?
3. If you were going to use the OERL Web site in teaching your courses, how exactly would you use it?
4. Do you consider the Professional Development modules appropriate for inclusion in your courses?
5. What feedback would you have on the auxiliary functions of the Web site?
6. Anything else?

Selection of Expert Panelists. The panelists were selected so that the site's usefulness could be explored from a variety of perspectives. Thus, the selected panelists had diverse evaluation specialties and methodologies.

The members of the expert panel were

- Dr. Christine Christie (Claremont Graduate School)
- Dr. David Fetterman (Stanford University)
- Dr. Nick Smith (Syracuse University)
- Dr. Robert Stake (University of Illinois at Urbana-Champaign, Emeritus)
- Dr. Mike Trevisan (Washington State University).

These faculty members teach a variety of evaluation or evaluation-related courses, including (1) introductory or procedural courses in evaluation; (2) evaluation theory; (3) advanced seminars or evaluation topics; (4) practicums in evaluation; (5) courses focusing on particular topics, such as evaluation for policy makers; and (6) methods courses, such as case study and qualitative methodology. In addition, each of these panelists either had experience in evaluating NSF-funded projects or had consulted with NSF on issues of evaluative practice.

Findings

The findings below are grouped by theme. They represent the combined feedback that panelists provided, based on their written reports, individual interviews, and conference calls.

Suitability of OERL as a Resource for Graduate-Level Evaluation Courses. The OERL team predicted that the suitability of OERL as a supplement to graduate evaluation courses would vary with the type of graduate-level courses and students. The responses of the panelists affirmed this prediction. Generally speaking, panelists reported that the depth and rigor of the site's resources were about right for at least some of the students some of the time. Appropriateness depended on how complex a treatment the students were prepared to address, and which aspects of evaluation the students needed to focus on at a particular time in their education. For example, one panelist said that the

professional development modules on Questionnaire Development would be appropriate for use in introductory courses, but the module on Methodological Approach and Sampling would be appropriate for more advanced courses covering outcomes-based approaches. OERL does not cover evaluation theory and thus was not considered useful in courses that focus largely on theory. A different panelist reported that with adaptation OERL could prove useful in introductory, practicum, and advanced evaluation courses.

Another component of OERL that was cited as being especially valuable to graduate students was the contact information about contributing projects. The information on contributors to OERL provides a network of evaluators that students can contact for further information about the artifacts that are posted on the site. A strength of OERL identified by one panelist was the use of examples from actual evaluations. This panelist maintained that an invaluable feature of professional growth for evaluators is learning through experience. Evaluators can also learn from one another: OERL provides information that evaluation practitioners can use to network with one another about types of interest.

Various panelists described ways they would use OERL in their practice-oriented courses. They would ask students to

- examine the examples of evaluation artifacts and features,
- critique the site's resources,
- critique the evaluation designs specified in a project,
- use the site's resources in writing a hypothetical proposal, and
- review auxiliary functions (alignment tables, glossary, criteria).

One panelist described how the OERL alignment tables could be useful to the students in his evaluation practicum. (The OERL alignment tables present quality criteria based on the *Program Evaluation Standards* (Joint Committee, 1994) that are appropriate for evaluation plans, instruments, and reports.) In the practicum, the panelist asks students to assess the extent to which their class evaluation project meets the *Program Evaluation Standards*. He asserted that OERL could be very helpful to students who are developing evaluation proposals, assuming sufficient lead time and faculty supervision.

The Ideal Web Site. In their reactions to OERL, the panelists revealed their views of what would constitute an ideal Web site for furthering the betterment of evaluation practice—visions that OERL in its current state did not completely fulfill. Their visions included a site that

- addresses evaluation theory,
- represents a greater diversity of evaluation approaches,
- stresses the importance of defining challenging issues for every project evaluation undertaken,
- addresses the field of evaluation in its rich complexity rather than simplifying the complexity to the degree present in OERL,
- presents a less prescriptive tone,
- emphasizes the many decision points that occur in an evaluation rather than making evaluation activities seem too mechanistic,
- stresses the role of the stakeholder in evaluation practice,
- stresses the usefulness of logic models, and
- presents a larger collection of plans and reports in their entirety.

The last point was made by one panelist who argued that there were not enough complete plans and reports in the OERL collection to call it a library. This criticism zeroed in on the dilemma that the OERL development team has always faced. Many of the artifacts submitted by evaluators to the OERL collection do not exemplify uniformly sound evaluation practice. The lack of consistently sound practice found in many of the submitted materials is a testament to the importance of a Web site for the professional development of evaluators, yet it also testifies to the limitations of relying heavily upon examples from actual practice for instructional purposes. For this reason, OERL includes professional development modules that rely, for pedagogical purposes, on made-up examples in addition to examples from actual evaluations.

Clarity and Purpose. Some panelists recommended the following changes to the site:

- Organize the resources in the collection primarily by the key features of an evaluation (e.g., goals, purpose, questions, context) rather than by project type (e.g., curriculum development, teacher education, faculty development).
- Add a site map that is aligned with a logic model to help individuals navigate the site in relation to the different phases of the evaluative activity.
- State the site’s focus on goal-driven, outcome-based approaches to evaluation on the home page.
- Use more visual representations to help users understand the professional development modules, alignment tables, and other auxiliary resources.

Professional Development Modules. Panelists recommended a number of changes to the professional development modules. Most felt that the presentation of content was appropriate, at least for beginning students. Recommendations for improvement included

- broadening the treatment of methodological approach,
- improving the user interface and making it “livelier”, and
- changing the tone to be less prescriptive.

Suggestions were made to include greater attention to theory, logic models, consulting with stakeholders, the importance of obtaining clearances from institutional review boards, and the requirements of the federal government’s Office of Management and Budget. Many of these topics were addressed in the additional professional development modules that have been created since the expert panel was conducted.

Auxiliary Functions. All resources on the site that are not part of the artifact collection (plans, reports, instruments) and the professional development modules are referred to as auxiliary functions. These include the user scenarios, contributor information, search tool, alignment tables, glossary, criteria, and FAQs.

The alignment tables contain (1) glossaries that define the components and features of plans, instruments, and reports; (2) quality criteria for plan writing, instrument writing, and report writing; and (3) alignments of the criteria to the *Program Evaluation*

Standards (Joint Committee, 1994). The criteria and glossaries can be accessed separately, or as part of the alignment tables. Most panelists commented that the various features of the alignment tables provided a useful framework. One suggested that the quality criteria might be expanded to include attention to alternative ways of reporting information. Another panelist wanted some acknowledgement that the “Joint Standards fail in such matters as recognizing the problems of over-promising in proposals.”

One offered qualified praise of the criteria and glossary but pointed out that OERL, with its outcomes-based scope, did not meet his vision of the breadth of approaches that should be addressed on an ideal site. Regarding the glossary feature in particular, comments varied widely. For example, one panelist commented that the glossary was unnecessary because the *Evaluation Thesaurus* (Scriven, 1997) is published and available. Another praised it as being useful and clearly organized.

Regarding the frequently asked questions (FAQs), several panelists suggested a reorganization in which the questions about evaluation were separated more clearly from questions about the Web site.

Generally, the other auxiliary features were praised as well presented and useful, though one panelist felt that there were too few user scenarios. (These scenarios provide hypothetical examples of how the site can be used.)

Discussion

Responses from panelists provided many useful structural and substantive suggestions that the OERL team could act on to improve the Web site. Responses also illustrated the challenge of developing a resource for the broader evaluation community, a community with varying beliefs, methodological preferences, and assumed roles for evaluation (Mark, Henry, & Julnes, 1999). Implicitly, responses also illustrated the difficulty of developing a single resource for graduate education for different faculty who hold differing ideas about teaching and learning as well as differing perspectives on evaluation, and who serve different student groups.

A key issue for graduate training of evaluators is how OERL should be used in class and what role instructors should play in enabling that use of the site. Some panelists evaluated the usefulness of OERL as a stand-alone tool that graduate students would use independently to complete class assignments. Others, however, evaluated OERL with the expectation that the instructor would provide guidance for students' use of the site.

Thus, it would be of value to develop a set of recommendations about the student group(s) for which OERL would be most appropriate, and how it might best be utilized for these students. In addition, the level and type of faculty support might be suggested. In short, an instructor's manual could be developed for OERL.

Another issue that came up during the discussion is the purpose for which OERL was created. Once the panelists learned that OERL was deliberately designed to address the goal-driven, outcome-based approach that many NSF evaluations adopt, debate occurred about how much this design limited OERL's usefulness. While most panelists agreed that the use of OERL to support NSF evaluations was defensible, there was disagreement about using OERL for broad training purposes.

Underlying many of the panelists' critiques of the OERL approach is a philosophical debate about how evaluation should be taught. Assuming that the site is appropriate for introductory-level graduate students, a question arises as to whether they are better off learning a fairly straightforward step-by-step, process-oriented approach at first—an approach that OERL presents—and later get immersed in the theories, complexities, and controversies in the field. Or, is it better to start them off with the breadth of evaluation approaches, then expose them to technical detail about individual approaches later? Another instructional issue is whether contextualized examples from actual evaluations should be heavily used or whether instructors should rely on hypothetical examples that can present quality practices in a “purer” form.

Current Development Work on OERL

Some topics for the professional development modules that were recommended by the panelists have been addressed in new professional development modules. These include modules about Instrument Selection and Triangulation, Selecting a Learning Assessment Instrument, Designing an Observation Instrument, Using Observation Instruments in the Field, Preparing an Interview Protocol, Administering Interviews, Data Collection Procedures and Schedule, and Presenting Results and Interpretations.

Other current development work on OERL includes reviewing and posting from the collection of NSF-funded evaluations of additional complete instruments, as well as complete versions of and excerpts from plans and reports. Also under development is a logic model and site map to assist evaluators in using the site more effectively.

References

- Joint Committee on Standards for Educational Evaluation (1994). *The program evaluation standards: How to assess evaluations of programs* (2nd ed.). Thousand Oaks: Sage.
- Mark, M. M., Henry, G. T., and Julnes, G. (1999). Toward an integrative framework for evaluation practice. *American Journal of Evaluation*, 20, 177–198.
- Scriven, M. (1991). *Evaluation thesaurus* (4th Ed.). Thousand Oaks: Sage.