

Editorial

Outcome-based evaluation

Michael Seadle

The author

Michael Seadle is Editor of *Library Hi Tech*.

Keywords

Evaluation, Digital libraries, Projects

Abstract

Outcome-based evaluation is increasingly used for digital library projects. The point of outcome-based evaluation is to establish a project's effectiveness with its target audience. Although many small-scale human-oriented project outcomes do not readily lend themselves to survey or statistical analysis, other research methods can be used to evaluate the amount and quality of a project's effect. Anthropology offers not only a suite of observational methods, but standards for persuasiveness that apply equally to academics and to legislators.

Electronic access

The Emerald Research Register for this journal is available at <http://www.emeraldinsight.com/researchregister>

The current issue and full text archive of this journal is available at <http://www.emeraldinsight.com/0737-8831.htm>

Introduction

I have just come from a presentation on outcome-based evaluation. This kind of evaluation is increasingly used (and sometimes required) for digital library projects. The temptation to treat evaluation as an afterthought particularly infects projects whose real goal is to produce a particular digital object, such as digital copies of a collection of books. Those objects are as concrete as anything can be in the digital world, and end-users seem ephemeral in comparison: anonymous, hard to find, hard to persuade to cooperate with response forms.

These end-users represent a real problem, since the point of outcome-based evaluation is to establish a project's effectiveness with its target audience. Reducing the data collection to something more casual than is required for peer-reviewed research journals is a bad option. The evaluation needs to persuade both the highly sophisticated agency (or foundation) staff who approve the project, and the people who ultimately provide the money. For many digital projects, this means the Congress, the Parliament, the *Bundestag*, or other equivalent legislative body.

The information standards for these bodies are high. Many members are lawyers accustomed to questioning evidence, and most have staff capable of analyzing any dubious data. Statistics from an ill-designed questionnaire could do a real disservice to the funding agency if they inspire ridicule. Claims that data show a statistically significant tendency ($p < 0.05$) could be completely unpersuasive, if an obvious selection bias vitiates a required assumption of randomness.

Although many small-scale human-oriented project outcomes do not readily lend themselves to survey or statistical analysis, other research methods can be used to evaluate the amount and quality of a project's effect. But such methods must be used with a care and caution that respects their standards of evidence, or the conclusions will have no more validity than a misapplied *t*-test.

Methodology

In my research I generally use anthropological theory and methods. They seem particularly well suited to understanding the impact of



projects on particular groups of people. Ethnography offers not only a suite of observational methods, but standards for persuasiveness that apply equally to academics and to legislators.

Ethnography is not a simple matter of asking a few questions or doing some observations. As Pitman and Dobbert (1986) explain:

Whereas [popular ethnography] relies on observation and interview data to produce a descriptive picture that is then used for evaluation purposes, ethnographic evaluation is founded on anthropological theory.

Fetterman (1986) goes further:

One of the most important attributes of a successful ethnographer is his or her ability to use a cultural perspective – in the artistic and literary Geertzian tradition – to decipher reality.

Geertz (1988, pp. 4-5) himself has written about the standards for persuasion in anthropological research:

The ability of anthropologists to get us to take what they say seriously has less to do with either a factual look or an air of conceptual elegance than it has with their capacity to convince us that what they say is a result of their having actually penetrated (or, if you prefer, been penetrated by) another form of life, of having, one way or another, truly “been there.” And that, persuading us that this offstage miracle has occurred, is where the writing comes in.

A few years later he continued to struggle to define the standards for proof:

The canons of anthropological “proof” being what they are (mimicries of sterner enterprises like mechanics or physiology) that is indeed how such doubts are most often phrased, and, to the degree they are, most often quieted. Footnotes help, verbatim texts help even more, details impress, numbers normally carry the day. But in Anthropology anyway they remain somehow ancillary: necessary of course, but insufficient, not quite to the point (Geertz, 1995).

The standards for anthropological evidence are not more lax than those for statistical hypothesis testing. They are more complex, harder to define, and ultimately rely on literary skills, breadth of knowledge, and skilled observation. In other words, they are not easy mechanical tasks that can be accomplished with little training in an afternoon, but they are designed particularly for understanding nuances of human behavior, and can readily be applied to evaluating small effects from small changes in the cultural environment.

Example

“Feeding America: the 19th Century American Cookbook Project” is an Institute of Museum and Library Services project being run at Michigan State University Libraries. The project plans to use fifth grade students as part of its evaluation process. The goals included:

- having students use the books “to understand the changing role of women in society, particularly the degree to which food technology freed them to take on roles outside the household”; and
- interesting the general public “in uncovering how their great-grandparents used to prepare and eat food” (Seadle and Berg, 2001).

The grant proposal talked about using focus groups with students and inviting the public, especially professional chefs, to volunteer for phone or e-mail interviews, as well as gathering statistics from the Web site. These represent important formal data-gathering processes, but an experienced anthropologist would recognize that their formality in part undercuts their value as representations of the cultural impact of the project.

Three unsolicited interactions may in fact show as much or more about the real effect of the project. The first came from a journalist for the *Detroit Free Press*. She read about the project from the prototype Web site, and approached us about doing an article that would describe the books and cooking implements, and would include modernized recipes tested in the paper’s kitchen. The newspaper represents a good measure of what a general reader will care about. The fact that a major urban newspaper would send reporters and photographers to create a multi-page story about the project shows the impact it had in even embryo form. The story also guaranteed early attention from subscribers.

The second interaction came through the university press, which showed strong interest in the possibility of establishing on-demand facsimile publication of the cookbooks. Like the *Free Press*, they anticipated a high level of public interest in the books and recipes. While on-demand publication required little investment, other digital books clearly excited them less. Their estimate was that a market for historic cookbooks exists. If people will actually

pay to own a copy, the project will have made a monetarily measurable impression on their lives.

The third interaction came from a colleague at another university who suggested digitizing some cookbooks in his own library's collection, and adding them to our site as a shared resource. He talked about potentially spending \$10,000 on the project. Budget woes due to the slow economy may delay or prevent carrying this out, but the interest showed another quantifiably positive effect on an end-user.

Three or four interested end-users is not a statistically significant sample, but the money they wanted to spend or expected to make, and their experience with general readers, imply a strong effect, especially in comparison with similar projects. Many of the basic digital publication techniques for Feeding America came from an earlier project called "Shaping the Values of Youth: A Nineteenth Century American Sunday School Book Collection," and yet that project garnered none of these kinds of interactions.

Conclusion

While these interactions seem like good news for the project, they have a troubling side as well. The interest seems entirely content-based. People like historic cookbooks. The idea of seeing them in facsimile form catches their imagination. Important added-value features such as the three-dimensional representations of cooking tools, the introductory essays, the author biographies, the glossary, and the search tools have evoked little more than polite nods. The reason could be that people take such features for granted in Web-based publications and would find the whole project less exciting without them. That is a comforting reflection with, however, no basis in evidence whatsoever. Perhaps the focus groups will provide an opportunity to probe further.

No conclusions from these interactions would be valid without some consideration of modern culture in the USA. Food has become a major interest in the past 40 or so years. Restaurants have grown in number and variety, as have specialty food stores, and cookbooks occupy shelf-space in all except the very smallest and most specialized bookstores. A project using

cookbooks will have an effect on the public because it feeds a well-established interest. A willingness to buy does not necessarily translate into a better understanding of the conditions of life in nineteenth century America, but bringing readers to passages about hints to young wives (Lea, 1869) or the dangers of skirts catching in fire in front of open hearths at least creates the opportunity for unplanned learning (Knowlton and Rice, 1831).

Anthropological observations do not guarantee positive or even particularly clear results. They do, however, offer a way to judge how small numbers of humans interact with small projects, when other more statistically-based theoretical models fail because of sample size or required assumptions.

Nonetheless, novice observers should be warned: for outcome-based evaluations to be useful, the evaluator must beware, or at least be aware, of private prejudices. Interpretations that do not take into account personal interests as well as broad cultural factors can be as invalid as a miscalculated statistic.

References

- Fetterman, D.M. (1986), "Evolution of a discipline", in Fetterman, D.M and Pitman, M.A. (Eds), *Educational Evaluation: Ethnography in Theory, Practice, and Politics*. Sage, Beverly Hills, CA, p. 216
- Geertz, C. (1988), *Works and Lives: The Anthropologist as Author*, Stanford University Press, Stanford, CA.
- Geertz, C. (1995), *After the Fact: Two Countries, Four Decades, One Anthropologist*, Harvard University Press, Cambridge, MA, p. 17.
- Knowlton and Rice (1831), *Cook Not Mad, Or, Rational Cookery: Being a Collection of Original and Selected Recipes*, Watertown, NY, p. 77, available at: <http://digital.lib.msu.edu/cookbooks/image.cfm?TitleNo=4&Image=079>
- Lea, E.E. (1869), *Domestic Cookery, Useful Recipes, and Hints to Young Housekeepers*, Cushings and Bailey, Baltimore, MD, p. 284, available at: <http://digital.lib.msu.edu/cookbooks/image.cfm?TitleNo=26&Image=288>
- Pitman, M.A. and Dobbert, M.L. (1986), "The use of explicit anthropological theory in educational evaluation: a case study", in Fetterman, D.M. and Pitman, M.A. (Eds), *Educational Evaluation: Ethnography in Theory, Practice, and Politics*, Sage, Beverly Hills, CA, p. 78.
- Seadle, M. and Berg, P. (2001), "Feeding America: the historic American cookbook project", grant proposal, available at: www.lib.msu.edu/seadle/Grant_proposals/Feeding_America.htm