

Cow One Is Not Cow Two

Harris H. Shettel
Evaluation Consultant
Rockville, MD

Steve Bitgood (pages 8-10 of this issue) has made a cogent and analytic argument in support of the need for the term "remedial evaluation" as one of the weapons in our exhibit evaluation armamentarium. At the risk of adding one too many commentaries, I would like to take a slightly different approach to the subject.

One of the most definitive ways of finding out if a particular verbal message is meaningful is to see if the recipient of the message acts in a manner consistent with the intended meaning. Assuming good intentions, if my wife asked me to get her a glass of water and I, in fact, got her a glass of water, one could say the message was clear and unambiguous. However, if she said, "I wanted ice in the water" I would say "Why didn't you tell me you wanted ice in the water?" The message in this case was only partially meaningful. The next time she asked for a glass of water, I might well begin a series of questions designed to clarify her true meaning - "Do you want ice in it?" "Do you want a big or little glass?" "Would you like a slice of lemon in it?" One might say that the number of questions one needs or ought to ask in response to a message (verbal or written) is negatively related to the perceived clarity of the message.

What, you may well ask, does this have to do with the debate over the use of the term "remedial?" Let us say that I was asked by a museum person who was in the process of preparing a grant request for a new exhibit, to write up a brief description of how I would go about doing a front-end study, a formative study, and a summative study of this exhibit. I would have relatively few questions to ask about the first two of these items. In fact, at this general level I have "boiler plate" descriptions of front-end and formative evaluations that have worked very well in countless proposals. However, if the person making this request happened to be Roger Miles, I would have a major question to ask him about what he meant by "summative evaluation." And this question would take the general form of "Are you only interested in finding out how well the exhibit met its overall objectives or are you interested in making changes to the exhibit based on the findings of the evaluation?"

Depending on the answer to this question, my proposal write-up would look quite different. If Roger only wanted a kind of "final report" on how well the main messages of the exhibition were communicated to visitors, I would talk about a relatively simple pre-post design, with questions aimed at

visitor understanding of these messages. However, if the intention is to make changes to the exhibition in those areas in which it is not doing well (both "within" and "between"), I would write about detailed tracking studies, both casual and cued visitor studies on each of the major components of the exhibition, and extensive analyses of the data so that the weak areas were not only pinpointed, but possible "fixes" would be recommended. I would go on to point out that these changes (which may range from "moving furniture around" to writing new labels) would have to be tested to see if they "worked," and if they did not, other possible solutions would be tried and tested. This iterative process would (ideally) continue until the exhibition was performing at its intended level of effectiveness.

It is even possible, as a result of this extensive and diagnostic data collection, that mock-ups of new exhibit elements would be prepared and tested (just like formative evaluation) or, in even more extreme cases, that visitors would be interviewed to find out what possible misconceptions they might have so that these misconceptions can be specifically addressed in the exhibition (just like front-end evaluation). I have no problem with this "looping back" to earlier stages of the exhibit development process as a result of doing a remedial study. And I would be comfortable in saying that I was doing front-end or formative work to help improve an existing exhibition. In this case "purpose" would take precedence over "when."

Depending on how comprehensive it was, it is also possible (even likely) that a person doing a Screven, Bitgood, Shettel kind of summative evaluation might come to the conclusion that the exhibition is a bomb, and they would probably have some pretty good ideas about ways in which it might be improved. He or she might then propose that while it was not the original intention of the study to change the exhibition, it appears to be highly desirable to do so. Perhaps more detailed visitor data are needed before the change process is initiated. A summative study has now "morphed" itself into a remedial study. I don't think that this requires any great intellectual wrenching of the mind to accommodate. Few categories in science are airtight (e.g., the discovery of hermaphrodites sure messed up the neat "male," "female" dichotomy!). I don't think, by the way, that one is likely to "morph" the other way - from remedial to summative.

A real-world example of the confusion that existed before remedial evaluation was given its own niche can be seen in the "Man in His Environment" study I did for the Field Museum in 1976. Since the exhibition was completed when I did the study it was called a "summative" evaluation. (I had earlier done two small formative studies during exhibit development.) While the "purpose" of the study was mainly to satisfy a requirement of the granting agency to find out if the completed exhibition was successful in communicating

its main messages to its visitors, I designed a fairly comprehensive data collection effort, one that, with one exception (cued testing), would satisfy the requirements of a remedial study. The final report noted a number of specific "problems" the exhibition had, but since there were no provisions in the budget to make any modifications to any of the existing exhibit elements, this study did not "morph." (*One small label* was added to a sculpture that we had already learned in the formative study was sending the wrong message to over 90% of visitors, but no testing was done to find out if the new label was effective!)

What I think the above examples show is that there is simply too much baggage that has attached itself to the single and exclusive term "summative." When activities as different as those I have described are subsumed under a single term, that term no longer meets the criterion of clarity or even of simplicity.

A field of linguistic analysis that was popular in the '40s and '50s, "General Semantics," led by Count Alfred Korzybski (1933), made a fetish of pointing out that things are often not what they seem even if we give them the same name. Their mantra was "Cow 1 is not Cow 2," reminding us that we blur important distinctions when we make our categories too big. (In fact, even "Cow 1" on Monday is not the same "Cow 1" on Tuesday!) Of course, such extreme, fine-grain discriminations can turn any discourse or communication into a hopeless mess (unless you are buying a cow), which is probably why "General Semantics" and the Count's magnum opus, *Science and Sanity*, have pretty much disappeared from the academic scene. Nevertheless, the opposite extreme, represented by the 14th century Franciscan, Occam, who is said to have said that "Entities are not to be multiplied without necessity," and "It is vain to do with more what can be done with fewer," (Franciscans take a vow of poverty!), seems to fuel the furnace of those for whom parsimony is to be preferred over clarity. Einstein said (if Roger can call the muse of Occam, I don't see why I can't do the same for Einstein), "Things should be made as simple as possible, but not too simple."

References

- Korzybski, A. (1933). *Science and Sanity*. The Science Press, Lancaster, PA.
- Shettel, H. (1976). *An Evaluation of Visitor Response to "Man in His Environment."* Available as Technical Report No. 90-10. Jacksonville, AL: Center for Social Design.

Suggested Readings on Evaluation Terminology and Methodology

The following articles should be of help to those who are unfamiliar with the classification debate discussed in this issue. The articles range from beginning to advanced readings.

O'Brien, M. (1992). What's Visitor Evaluation All About? *Visitor Behavior*, 7(2), 5-10.

This is a well-written, readable article that gives an overview of the four types of evaluation (front-end, formative, remedial, and summative) described by Screven (1990) and Bitgood and Shettel (1994).

Korn, R. (1994). Studying Your Visitors: Where To Begin. *History News*, (March/April), 23-26.

Another easy-to-read introduction to evaluation. This article, consistent with Miles (1993), categorizes evaluation into three types (front-end, formative, and summative).

Screven, C. G. (1990). Uses of Evaluation, Before, During, and After Exhibit Design. *ILVS Review*, 1(2), 36-66.

This is the most ambitious and detailed article on exhibit evaluation. Screven gives many examples and describes detailed methods for carrying out exhibit evaluation. The term "remedial evaluation" was first introduced in this article. While this article is "heavy" reading, it is well worth the time and effort.

Miles, R. (1993). Grasping the Greased Pig: Evaluation of Educational Exhibits. In S. Bicknell and G. Farmelo (Eds.), *Museum Visitor Studies in the 90s*. London: Science Museum of London. Pp. 24-33.

Miles presents an elegant and cogent defense of the philosophy and methodology of visitor studies. In addition, he challenges the need for the concept of "remedial evaluation" arguing that if evaluation is classified by when it occurs and what is being evaluated, only three types of evaluation are necessary.

Bitgood, S. & Shettel, H. (1994). The Classification of Exhibit Evaluation: A Rationale for Remedial Evaluation. *Visitor Behavior*, 9(1), 4-8.

This article defends the use of remedial evaluation and the importance of "purpose" as a criteria for distinguishing among evaluation types.