Demystifying evaluation: a brief guide to the evaluation of interpretive media, activities and programmes

Dr. Philip Ryland and Dr. Sunita Welch

AHI Best Practice Guidelines 12. August 2016.

So, what is ‘evaluation’?

Evaluation is a systematic process of determining something’s value, worth or merit. When you evaluate your interpretation it will help you to understand whether it is meeting its objectives or not and will therefore help you to develop the specific media or entire programme further. It will also help you answer any ‘so what?’ questions about your work and therefore help you to explain what is different because of your interpretive project.

Evaluation, as a process, can be simplified in terms of a number of key questions, including; ***Why? When? What?*** and ***How?*** – These questions will be briefly explored in this guide.

Evaluation stage 1: *Why?*

Evaluation plays a key role in improving the visitors’ experience of your interpretation and can also play a significant role in helping you to gain external funding. It can be used to test your initial interpretive ideas and later to determine the effectiveness of the overall experience and whether it is meeting its objectives as well as the value of individual media or programme(s) within the overall experience. You can also use the results of any evaluation to assess the value of the interpretation to the site as a whole as well as to inform site-based policy and planning decisions.

Table 1. Why evaluate?
To determine the educational and recreational impact of your interpretation, in both the short and longer-term
To assess the cost effectiveness of the various interpretive media used
To assess the contribution of your interpretive programme(s) or, individual media, to the visitors overall on-site experience

To assess the contribution of your interpretive programme(s) or, individual media, to the modification of visitors overall on-site behaviour
To provide accountability
To convince managing authorities, public bodies and others of the value of your interpretation, providing evidence that your interpretation addresses and achieves important public and site-based goals
To help inform policy and planning decisions
To help provide the public with a way to indicate their response to interpretive services, beyond simple attendance data
To assess the contribution of the overall interpretive programme(s) to achieving sustainable goals, in terms of the management and perpetuation of the site and its resources

(Table based on the work of Diamond *et al*., 2009; Ham, 1992; Knudson *et al*., 1995; Veverka, 1994; Ward and Wilkinson, 2006).

Thus, it is suggested that in evaluating your interpretation, the aim and objectives of your organisation can be more fully achieved, equally site managers can be better informed which can in turn lead to the development of a higher quality of interpretation provision in the future.

Evaluation stage 2: *When?*

Evaluation should be an on-going process and thus it should be an integral part of the regular review of your on-site interpretation. There are a number of ways to divide the stages in the evaluation process, typically however there are five forms of evaluation which can be used to support your interpretation and these are; ***front-end, formative, remedial (process), summative (outcome)*** and ***impact*** evaluation.

Front-end evaluation is undertaken at the earliest stage of the interpretive planning process and typically involves market research focusing upon visitor knowledge of the site as well as their levels of interest in potentially new interpretive themes or stories. If you are looking to attract new audiences, you will need to locate them off-site and work with them in their community. You might also explore the site-based resources visitors are most interested in viewing, who they come with, how long they stay and how they wish to engage with the site and its resources. The results from your front-end evaluation should help you to tailor your interpretation more effectively to the needs of your visitors.

Formative evaluation typically occurs during the implementation phase to test a 'trial' version of the media or programme being developed. In an exhibition, formative evaluation might be conducted on prototypes, mock-ups or on inexpensive props before the final media are produced. You might specifically test visitor reaction to the media, such as their attention or understanding as well as their recall of, or attitude to, the messages it is trying to communicate. For interactive exhibits, you might test all the moving parts, what works and what might need to be changed. Whilst for an interpreter, formative evaluation might be used to help them to establish the correct level at which to pitch an interpretive communication with a 'test' audience.

Process (remedial) evaluation is the on-going monitoring and assessment of how your interpretation is working. It is often used to test and therefore build specific relationships with your target audience(s). It is also used to check that all the elements of the media or programme work successfully together such as the sound, lighting, graphics etc.

Summative (outcome) evaluation is generally the most widely and regularly used form, it is carried out after the interpretive media or programme has been completed and is most often used to assess its success in relation to its original objectives. However, interpreters should be aware that the biggest weakness of relying solely on summative evaluation is that you often cannot go back and change the interpretation as a result of your increased knowledge. In summative evaluation, visitors are typically encouraged to tell staff what they think about their experience often through a questionnaire, interview or focus group. The direct and unobtrusive observation of visitors as they view or undertake the media or programme is another common method of undertaking summative evaluation.

Summative evaluation can also include a time-frame element where perhaps a questionnaire or telephone interview is conducted some weeks after the visitors have returned home, typically to assess longer-term learning and/or recollection of experiences, this is known as post-occupancy (POE) evaluation.

Impact evaluation is the 'pot of gold' at the end of the rainbow! Many funding bodies and organisations want to know the long-term impact of interpretation – the '*so what?*' of your site. Are your visitors more engaged as a result? Do they take away ideas that impact upon their daily lives? Has any longer-term behaviour change occurred?
If you have undertaken good front-end evaluation and allowed the time and resource to re-valuate your media or programme on a regular basis, then the outcomes from this form of evaluation can be very powerful.

Finally, it is worth considering that your evaluation can be undertaken as 'goal-referenced' or 'goal-free' where 'goal-referenced' relates to you measuring specific objectives whilst 'goal-free' allows you to undertake a more open evaluation which may reveal unexpected effects or impacts.

Evaluation stage 3: *What?*

Typically, an evaluation of your interpretive media or programme(s) will explore one or more of the following four categories:

- evaluating the ***visitors*** (existing and potential visitors) to examine their interest in, and reaction to, the interpretation on offer;
- evaluating the ***whole programme*** to identify which media have consistently achieved their objectives and to modify others where greater effectiveness might be achieved;
- evaluating the ***performance of interpreters*** and helping them to identify ways in which they could improve their delivery;
- evaluating the ***overall productivity*** of the programme and its associated facilities to determine whether money and effort are being managed effectively on the site.

Evaluation stage 4: *How?*

The range of methods you adopt typically might include questionnaires and interviews, focus groups, the direct participation by an 'experienced' observer, the unobtrusive observational studies of visitors as well as behavioural and mapping exercises. Modern technology can now also be used to support some of the formerly more intrusive methods, such as notably visitor observation.

Four broad categories of investigation have traditionally been identified, these include:

- the informal appraisal of a site's interpretation usually by a professional interpreter or other 'expert';
- the observation of visitor behaviour without their knowledge. Such techniques include the noting of routes taken around an exhibition, the amount of time spent looking at individual displays ('dwell time') and at which elements of each display, the levels and types of interaction with each display;
- techniques involving informal but direct contact with visitors for example, the use of self-completed questionnaires or short interviews as well as GPS tracking on larger sites;
- more formal contact with visitors often involving the use of a detailed questionnaire or interview or, their participation in a focus group exploring one or more topics in depth.

Table 2. The How of evaluation

Evaluation objectives	Type of information required	Evaluation technique	Desired outcomes
1. The visitor 1A) Their openness to interpretation 1B) Their attention to the interpretation	1A) Comparison between numbers engaging and the numbers on the site, as a whole 1B) Comparison between time visitors spend looking at the interpretation and the minimum time needed for exposure to the full message	For both; 1. Observations (track, time, watch, listen) 2. Questionnaires 3. Focus groups	For both; 1. Improvement in quantity and design of media 2. Improvement in the visual impact of media 3. Improvement in the appropriateness of media content, to the visitor and their needs
2. The impact of the media 2A) On the visitors enjoyment	Comparison of visitors enjoyment and satisfaction before and after the visit	1. Observations 2. Interviews 3. Questionnaire 4. Focus groups 5. Personal comments (verbal or written)	1. Improvement in visitor satisfaction and enjoyment with the site and its interpretive media.
2B) On the visitors learning	Comparison of visitors understanding of site-based topics interpreted, before and after the visit	1. Questionnaire (POE) 2. Interviews (POE) 3. Focus groups	1. Improvement in visitor knowledge of the site and their broader awareness of its fragility, importance, uniqueness etc.
2D) On the visitors behaviour	Changes in observed visitor behaviour	1. Questionnaire 2. Interviews 3. Observations 4. Focus groups	1. Improvement in visitor on-site behaviour towards the site and its artefacts, wildlife, people etc.

3. The <u>performance</u> of the interpreter	Evaluation of the interpreter's performance by their; ~ supervisor, ~ peers, ~ outside experts, ~ self-evaluation with the audience.	Individual or group critique, by; 1. Observations 2. Interview 3. Focus groups 4. Personal comments (verbal or written)	1. Improvement in the presentation of interpretive media by the interpreter. 2. Improvement in the interpreter's overall performance.
4. For <u>site managers</u> <u>Cost</u> effectiveness of the interpretation	4.1) In relation to capital, running, staffing and maintenance costs. 4.2) In relation to visitor numbers in attendance on the programme(s).	1. Record keeping 2. Observations 3. Interview 4. Questionnaire 5. Focus groups	1. Recognition of the cost-effectiveness pf the provisioning of interpretive media.

(Table based on the work of Diamond *et al*., 2009; Ham, 1992; Knudson *et al*., 1995; Merriman and Brochu, 2005; Moscardo, 1999; Veverka, 1994; Ward and Wilkinson, 2006).

Data collection methods

In summary, the five most usual methods of data collection undertaken are; Surveys, Interviews, Focus Groups, gathering Personal Comments and Observations.

Surveys (or questionnaires)

A series of open and closed questions which are read by the participant and their responses are either self-administered or are recorded by someone else. Self-administered is the least expensive method but often results in a limited number of responses. Surveys should be very clearly written, avoid jargon and technical terms and should not normally take longer than 5-8 minutes to complete, this equates typically to questions on a maximum of 2-3 sides of A4 paper. Surveys limited to one side of A4 paper tend to attract the greatest response rate. In terms of sample size, you should consider approaching 100-150 people when undertaking a survey.

Interviews

A series of open and closed questions administered by an interviewer who records the participant's responses. Interviews normally last between 15-45 minutes in length. You should ensure that the language of your interview is suitable for your intended interviewees. For interviews, the sample size largely depends upon the length of your interview and the nature of the topic being explored, but typically 5-15 interviews should be conducted.

Focus Groups

A meeting of 6-12 participants facilitated by someone who guides the group through a series of structured open questions, activities and/or discussions, it should not be merely a question and answer session. Focus groups also do not need to move towards a final consensus view! Focus groups can normally last up to 120 minutes, beyond 60 minutes you should include a break for refreshments.

The purpose of a focus group is to explore the differences in views between participants and to therefore help to understand the factors that influence their opinions, behaviours and/or motivations about the site and its resources. Focus groups can also be used to pilot new media, develop or test survey / interview questions or to explore data which has already been collected through observation, a survey or interview.

Gathering Personal Comments

Personal comments may be made by a single person based upon their on-site experience or may be the collective views of a group of friends or family members. Depending upon the nature of the site these personal comments may be captured in a whole variety of ways. Traditionally this would include a 'visitor's book' or 'comments card' which encourage visitors to write a short comment upon their experience either at a specific location or activity, or having been through an exhibition or, on the site as a whole. Today, personal comments are also likely to be widely reported on social media (such as Facebook or Twitter pages) but also through on-line feedback sites such as TripAdvisor. Comments of course may also be verbal where they are received by a member of on-site staff who can then collate them and convey them to the relevant member of staff.

Observations ('behavioural mapping' / 'value judgements')

Unobtrusive observation is used to measure the 'attraction power' as well as the 'holding power' of a media, activity or entire programme. No personal data (beyond broad demographics) is collected as it is nature of engagement rather than the individual participant which is of interest. Observations can be used to record routes taken around an exhibition or entire site ('tracking'), the amount of time spent looking at an individual display ('dwell time') and most typically, the types and levels of interaction with each display, this latter type of on-site observation is often referred to as 'behavioural mapping' but also includes 'value judgments' as the participant is making choices about what they undertake on site. A coding sheet is normally used to record the types of interaction undertaken by a participant as well as how long they interact with the media for. Listening to conversation at a specific exhibit or during an entire exhibition may also form part of unobtrusive observation where individual comments made, the nature of discussion within a group or the questions being asked are recorded.

All of these methods may be used in isolation but most often they are used in conjunction when evaluating a programme, media, exhibition or entire site.

Suggestions for further reading

Diamond, J., Luke, J. J. and Uttal, D. H. 2009. *Practical evaluation guide: tools for museums and other informal educational settings.* (2nd ed.). Lanham, Maryland: AltaMira Press. Part II – Evaluation Tools.

Jacobson, S. K. 2009. *Communication skills for conservation professionals* (2nd ed.). Washington, DC: Island Press.

Kellogg Foundation. 2006. *Logic model development guide*. Available from: www.wkkf.org/resource-directory/resource/2006/02/wk-kellogg-foundation-logic-model-development-guide

Moscardo, G. 1999. *Making visitors mindful: principles for creating quality sustainable visitor experiences through effective communication.*
Champaign, IL: Sagamore Publishing. Ch. 7. Getting to know your visitors.

Post, M. 2012. *Fearless evaluation*. A webinair for the US National Park Service. Available from: www.nps.gov/hfc/services/interp/interpPlanning/FearlessEvaluationManual.pdf

Sharpe, G. W. (ed.). 1976. *Interpreting the environment*. New York: John Wiley & Sons. Ch. 26. Research in Interpretation.

Ward, C. W. and Wilkinson, A. E. 2006. *Conducting meaningful interpretation: a field guide for success.* Golden, CO: Fulcrum Publishing. Ch. 13. Practising defensible interpretation: Evaluation.

Webb, E. J., Campbell, D. T., Schwartz, R. D. and Sechrest, L. 1999. *Unobtrusive measures*. Champaign, IL: Sage Publications.

Text by:
Dr. Philip Ryland, MAHI.
Associate Dean (Student Experience), Faculty of Management, Bournemouth University.

Dr. Sunita Welch, MAHI.
Education and Interpretation Manager, Brecon Beacons National Park Authority.

AHI Best Practice Guidelines 12.
Association for Heritage Interpretation, 54 Balmoral Road, Gillingham, Kent. ME7 4PG.
+44 (0)560 274 7737
admin@ahi.org.uk
www.ahi.org.uk

August 2016.