

Evaluating impact and quality of experience in the 21st century: using technology to narrow the gap between science communication research and practice

Eric A. Jensen

Abstract

Access to high quality evaluation results is essential for science communicators to identify negative patterns of audience response and improve outcomes. However, there are many good reasons why robust evaluation linked is not routinely conducted and linked to science communication practice. This essay begins by identifying some of the common challenges that explain this gap between evaluation evidence and practice. Automating evaluation processes through new technologies is then explicated as one solution to these challenges, capable of yielding accurate real-time results that can directly feed into practice. Automating evaluation through smartphone and web apps tied to open source analysis tools can deliver on-going evaluation insights without the expense of regularly employing external consultants or hiring evaluation experts in-house. While such automation does not address all evaluation needs, it can save resources and equip science communicators with the information they need to continually enhance practice for the benefit of their audiences.

Keywords

Informal learning; Popularization of science and technology; Public perception of science and technology

On-going evaluation tied to real-time results can enable science communication organizations to develop activities that stand a stronger chance of yielding positive impacts. Given the logistical barriers facing science communication organizations seeking to develop high quality evaluation (discussed below), such technology can be an invaluable tool for enabling science communication organizations to keep a finger on the pulse of their audiences. Many science communication practitioners would acknowledge the value of impact evaluation for assessing whether mission-related outcomes are being achieved. However, the practical barriers of required expertise, time and resources can make impact evaluation — particularly on an on-going basis — seem like an impossible task. Yet recent technological improvements have created new means of gathering and analyzing on-going quantitative and qualitative survey-based evaluation using automation (e.g. see qualia.org.uk or artory.co.uk). While social scientific expertise is required at the stage of designing the evaluation and survey questions, an automated system can run indefinitely providing insights to the organization for years without the need

for external consultants or in-house evaluation staff. This essay explores the reasons for needing such a technology-enhanced approach, as well as the strengths and limitations of different technology enhanced options for survey-based evaluations.

**Common
challenges in
science
communication
evaluation**

Given the complexity of science communication interactions — bringing together multiple individuals' values, assumptions, worldviews and meaning-making processes — the remarkable scenario is when positive outcomes are achieved. There is a great deal of existing knowledge in social science domains such as psychology, social anthropology, communication and sociology about the principles most likely to yield positive outcomes. Once these principles are put into practice through particular initiatives, robust impact evaluation can reveal the effective science communication approaches that should be extended or the ineffective methods that need to be stopped or adjusted. Of course, the professional field of science communication does not have a shared goal, or even a universally shared notion of what constitutes a 'positive outcome'. Moreover, 'effectiveness in this domain is not an obvious, unidimensional and objective quality (such as speed or distance) that can be easily identified, described, and then measured' [Rowe and Frewer, 2004, p. 517]. As such, the task of focusing and refining science communication goals is always a key challenge when developing an impact evaluation. However, this challenge is outside the scope of the present essay. Here, I focus on how to conduct robust evaluations with limited resources once those goals have already been refined.

It has long struck me as ironic that informal learning institutions touting the value of the scientific method for society can be so reluctant to employ a robust evidence-based approach to their own practices. However, I have come to see that there are good reasons for the paucity of impact evaluation in science communication organizations.

1. It takes time to learn how to design and conduct high quality data collection and analysis. Many informal learning and engagement professionals are just too pressed with other priorities to undergo such extensive training.
2. Science communication organizations do not recruit staff with social scientific methodological expertise within the institution.
3. There is a tendency for science communication organizations to draw upon external consultants to plug the gap in in-house knowledge and capabilities to conduct evaluation. In practice, this tactic often fails to produce quality evaluation research for these organizations because many of these consultants also lack appropriate social scientific training and methodological expertise. As such, they often produce spectacularly poor quality evaluations [e.g. see Jensen, 2014].
4. In lieu of robust evaluation, an organization sometimes conducts an anecdote gathering exercise focused on eliciting positive accounts of how wonderful a program is. This is not evaluation in any meaningful sense, however this kind of advocacy or campaigning activity is sometimes conflated with evaluation.
5. In particular, surveys and associated evaluation procedures are often flawed, with basic errors in survey design and sampling, compounded by limitations in data analysis and interpretation [e.g. see Jensen, 2015b].

6. Qualitative evaluation and research takes extensive training to conduct in an effective manner. From how to craft interview questions to how to document the context and conduct qualitative data analysis, qualitative research skills draw on an extensive body of methodological literature. This literature also clarifies the distinction between 'good' and 'bad' qualitative research [e.g. Thorne, 1997; Gaskell and Bauer, 2000]. When conducted effectively, qualitative evaluation can offer a depth of insight that is not feasible with quantitative methods, but a paucity of social scientific research training in the science communication domain has made this type of evaluation relatively rare [cf. for a positive example Dawson, 2014].

Measuring impact has moved up the agendas of many science communication organisations in recent years. However, it generally requires a commitment of resources to gather cultural impact data, as well as valid feedback data, with many organisations struggling to develop a robust evaluation framework. The fact that scarce resources are often invested in arriving at such unsatisfactory results is particularly problematic [Jensen, 2014]. These concerns were a key motivation underpinning the Qualia project (qualia.org.uk), funded by the UK's National Endowment for Science, Technology and the Arts, the results of which are discussed below.

Qualia took on the task of assessing the possibilities for using new technologies to replace or supplement existing evaluation methods. It is able to collect and analyze individual feedback, while automatically aggregating results in real time to look for patterns. The goal is to bring evaluation evidence and practice closer together.

Overcoming common challenges by automating evaluation

The Qualia project (qualia.org.uk) explored a range of automated quantitative and qualitative options for evaluation, including automatic smile detection, automated sentiment analysis of audiences' social media responses to events and automatic tracking of audience flows integrated into a public engagement festival smartphone app. The project involved critically evaluating the utility and validity of using smartphone apps and other tools for evaluation [Danielson et al., 2015]. The methods developed through this project allow digital data to be easily and inexpensively used by cultural institutions (including science communication organizations) to assess audience members' immediate responses to events in real time, as well as tracking audience responses over a longer time frame (e.g. over the course of a festival or longitudinally over a number of months or years).

Beyond concerns about accuracy, the major goal underpinning the search for technology-enhanced evaluation solutions is enabling public engagement organizations to have their own systems of evaluation that are not dependent on external consultancy. This type of 'evaluation enablement' means that a science communication organization has on-going access to useful evaluation evidence that does not require the active involvement of a third party, or additional incremental costs for data collection once a system is established. This essay therefore focuses on evaluation enablement for science communication practitioners through automation. To maintain a tight focus in this article, I will limit my discussion to survey-based options for automated evaluation, including (1) smartphone app-based evaluation and (2) email/online survey systems that do not require smartphone ownership for participation.

Establishing such automated evaluation systems based on surveys can enhance efficiency and quality for science communication organizations. Advantages of such technology-enhanced evaluation systems include:

- Less expensive than a market research company
- Far less staff time needed when compared to in-house data collection (effectively no on-going staff time needed)
- Better quality data available than is likely from either a market research company or in-house (in most cases)
- More extensive and timely data can be gathered than would otherwise be feasible
- Real-time automation of analysis provides results as the data rolls in. This means that science communication organizations can act on evaluation findings immediately, rather than having to wait for an end-of-project or annual report.
- Minimize logistical challenges to organize for data collection
- Organizations own and maintain full and permanent access to their data

In addition, some types of evaluation are only logistically feasible through automation, such as the seamless inclusion of multi-lingual data collection and gathering feedback on specific events within a larger experience such as a festival.

Automating science communication evaluation using a smartphone app system

A smartphone app-based system can be a valuable option for gathering feedback from audiences while they are still in close proximity to the event or activity being evaluated. In the case of the Qualia project, a prototype smartphone app was developed for audiences to use to improve their experience at the Cheltenham Science Festival, a major UK science communication event including numerous specific science communication activities for public audiences. This app was implemented, for example by having a scheduler feature to add events with automatic reminders, etc. [Danielson et al., 2015]. A micro-survey designed to gather feedback about the experience through four short questions (e.g. 'enjoyable?') was integrated into the smartphone app. An automatic feedback request was pushed to users when someone indicated they were attending an event. Questions were customized to organizations' specific requirements and event types. This kind of integrated evaluation within smartphone apps can also be added to an existing audience app.

A recent follow-on project from this prototype smartphone app (linked to open source database and analysis software) is Artory, a collaborative research and development initiative by ten cultural organizations in Plymouth, UK (artory.co.uk). This system uses a city-wide 'what's on' smartphone app listing of arts and culture events and activities as the starting point for measuring the quality of cultural experiences. Participation in providing feedback data is incentivized. Users are rewarded with 'Art Miles' for checking in at different venues and

providing feedback: the points that are accrued in this way can then be redeemed at participating venues for a range of offers, from coffee or tea to discounted tickets and special 'VIP' access to activities. However, participation is fully voluntary.

The measures used in this smartphone app were developed through a consultation process with all of the participating organizations [see Jensen, 2015a]. To avoid over-taxing app users / respondents, three different levels of feedback are employed with audience members invited to opt into further levels of depth in providing feedback in order to unlock greater incentives. The most basic level includes a matching pre- and post-visit survey item evaluating expectations and outcomes. The pre-visit question is as follows: 'What are you hoping to gain from attending this event? [tick all that apply]'. The paired post-visit question asks, 'What do you feel you gained from this event? [tick all that apply]'. The response options for both of these question are the same (Table 1) to make visible the differences between expectations and self-reported outcomes.

Table 1. Matching Response Options for Expectations and Outcomes in Artory app.

Fun Day Out
Learning
Cultural Experience
Community Experience
Inspiration
Family Time
Good Time with Friends
Professional Networking
Entertainment
Other

The truncated format of the response options is designed for use on a smartphone app-based micro-survey, where long response options can be awkward for respondents to use [Danielson et al., 2015].

This type of evaluation data is particularly valuable because it employs exactly the same phrasing and survey structure to gather feedback about public engagement experiences from multiple institutions. This enables measurement over time and across different cultural experiences in the same city, from live theater to galleries to museums and other settings. Indeed, this system enables *tracking of unique visitors* across different engagement experiences, enabling a new level of knowledge about how specific cultural experiences fit within a wider range of experiences for an individual

As an added practical benefit, respondents/users only have to enter profile information once instead of reentering their demographic data repeatedly over time and across different sites. Following the individual rather than the site or organization itself is a significant step forward in terms of audience studies and understanding audiences' relationships with cultural institutions. Using a smartphone app-based system is clearly an efficient way of achieving this step forward [also see Jensen, 2015a].

Automating impact evaluation using a web-based survey system

Unless a sophisticated experimental design is being used, accurately measuring science communication impact requires (at minimum) direct measurement of visitors' thinking or attitudes before and after the informal learning or engagement experience. Building on the know-how developed through the Qualia project, science communication organizations can establish a framework for on-going 365 day per year impact evaluation. When audiences register for an event in advance, they can be automatically invited to provide answer pre-event survey questions at that point. When audiences arrive at a site without prior booking, they can be asked to complete a short pre-visit survey on a tablet or iPad linked to the automated evaluation system. Either way, respondents can be automatically sent the post-visit survey after they return home from the science communication experience. Statistical analysis linked to graphs are then automatically applied to the data as it comes in, yielding real-time results that practitioners can view at any time. This is currently in place for a zoo in Brazil, which has no research staff but is able to view results from its automated visitor survey to learn how Spanish-, Portuguese- and English-speaking visitors are responding to the experience on a practical level, as well as evaluating their attitudes about wildlife and conservation. Making real-time research insights about audience engagement available to science communication organizations can create a stronger bridge between evaluation and practice allowing organizations to act immediately when problems are revealed.

Limitations of technology-enhanced evaluation tools

Access to digital tools, especially smartphones, can present barriers within particular demographic categories, especially along age and social class lines. Each digital technology brings its own patterns of participation/exclusion. For example, Norris [2001] highlights three distinctive forms of digital divide: 'global' (between rich and poor nations), 'social' (inequality within a nation) and 'democratic' (between those who use digital technology for civic, or 'public' purposes and those who do not). Given the tendency for science communication institutions, like other cultural institutions, to disproportionately serve those who are already economically advantaged and highly educated [Dawson and Jensen, 2011; Jensen, Dawson and Falk, 2011], the digital divide is an important issue for evaluation to attend to.

Two considerations weigh in favor of using digital technology for evaluation, despite the digital divide. First, on the whole, I would argue that it is in the interest of (potential) audience members from under-represented minorities, excluded and culturally deprived communities, to have science communications taking a more evidence-based approach with on-going evaluation integrated into their public experiences. This evidence-based approach can be enabled through digital technology. Moreover, organizations with a clear pattern of under-representing minority groups, should initiate further in-depth qualitative research to supplement an automated system and identify where improvements can be made to enhance social inclusion.

Secondly, this important concern pertaining to the digital divide is partly mitigated with an e-mail/online system, as these are among the more ubiquitous digital

technologies (along with cellular text messaging). Moreover, some features of such an automated system can enhance the social inclusion of audience research, most notably by opening up the ability to conduct multi-lingual audience research as a seamless and low-cost dimension of an automated system. Minority language speakers have long had their voices excluded from audience research, therefore a means of routinely including this audience group certainly contributes to an inclusion agenda.

In sum, we must be attentive to the risk of merely amplifying the voices of people from demographic categories already over-represented within a visitor population at a science communication institution through over-reliance to technology-enhanced evaluation. Of course, low-technology methods can also be exclusionary. Therefore, this risk needs to be carefully weighed each time a new evaluation system is set up. Key issues to consider include the current audience profile, alternative methods of gathering feedback or evaluating impact for those without the access to technology and the need to periodically undertake robust qualitative research to dig beneath the surface of an issue.

Conclusion

Communication is rife with potential for mutual misunderstanding, and science communication is no exception. Understanding longstanding principles of good communication [e.g. Spitzberg, 1983] and applying them to science communication practice is essential. However, at some point science communicators should go further and seek to robustly evaluate audience impacts to see if the desired outcomes are developing and why. Yet, this is rarely done, even within the most well-funded science communication institutions.

There are many possible reasons for the poor performance in science communication impact evaluation to date [Jensen, 2014; Jensen, 2015b], including a lack of appropriate social scientific research methods training amongst the ranks of science communication evaluation consultants. Automated evaluation system offer numerous advantages over existing options for science communication organizations seeking to gather audience feedback and demographic data. These advantages include lower costs, higher data quality and more extensive, individual level data about audience members and their experiences.

This essay has discussed how the use of robust open source technology provides a sustainable, adaptable, customizable and extensible vehicle for bringing science communication practice, evaluation and research closer together. It is even possible to make available 'in the moment' or real time data sets during and immediately after science communication experiences through low-cost technologies. Data harvesting, aggregation and analytic systems initially established for Qualia can be used to aggregate primary data and conduct deep levels of quantitative evaluation analysis.

In this essay, I have argued that automated systems can make real-time research insights about audience engagement available to science communication institutions that would otherwise be unable to afford high quality audience feedback and impact measurement. Gaining a robust understanding of audience experiences can enable science communication organizations to be better attuned to, anticipate and predict the changing needs and interests of current and potential

audiences, thereby delivering more effective experiences. The automation inherent in open source and data driven systems like Qualia and Artory is designed to open up audience analysis to a wide range of public engagement organizations, not just the select few that can afford expensive consultants or hire in-house expert technical staff. Meanwhile, well-funded institutions can also benefit from the high quality integrated data collection, with automated analytics tailored to the needs of the organization and its audiences. Moreover, automating routine survey data collection and analysis could free up resources for more in-depth qualitative research that it is not feasible to automate. In sum, I contend that the future of science communication should involve more extensive engagement with technologies to enhance the role of evaluation in enhancing practice. At the same time, we must ensure that this future enhances, and does not diminish, social inclusion amongst potential and actual science communication audiences. This can be achieved through careful evaluation design, which may include qualitative evaluation to thoroughly address the range of evaluation needs

References

- Danielson, K., Jenkins, J., Phillips, M. and Jensen, E. (2015). *Cheltenham Festivals: Real-time event feedback*. Report published online by NESTA. URL: <http://artsdigitalrnd.org.uk/wp-content/uploads/2013/06/Cheltenham-Festivals-RD-Project-Report.pdf> (visited on 20th August 2015).
- Dawson, E. (2014). “‘Not designed for us’”: How science museums and science centers socially exclude low-income, minority ethnic groups’. *Science Education* 98, pp. 981–1008.
- Dawson, E. and Jensen, E. (2011). ‘Towards a ‘contextual turn’ in visitor research: Evaluating visitor segmentation and identity-related motivations’. *Visitor Studies* 14 (2), pp. 127–140.
- Gaskell, G. and Bauer, M. W. (2000). ‘Towards public accountability: Beyond sampling, reliability and validity’. In: *Qualitative researching with text, image and sound*. Ed. by M. W. Bauer and G. Gaskell. London, U.K.: Sage, pp. 336–350.
- Jensen, E. (2014). ‘The problems with science communication evaluation’. *JCOM* 13 (01), C04. URL: http://jcom.sissa.it/archive/13/01/JCOM_1301_2014_C04 (visited on 20th August 2015).
- (2015a). *Enabling city-wide audience evaluation: Plymouth Artory metrics development*. Report published online. URL: https://www.researchgate.net/publication/281109028_Enabling_City-wide_Audience_Evaluation_Plymouth_Artory_Metrics_Development (visited on 20th August 2015).
- (2015b). ‘Highlighting the value of impact evaluation: Enhancing informal science learning and public engagement theory and practice’. *JCOM* 14 (03), Y05. URL: http://jcom.sissa.it/archive/14/03/JCOM_1403_2015_Y05.
- Jensen, E., Dawson, E. and Falk, J. (2011). ‘Dialogue and synthesis: Developing consensus in visitor research methodology’. *Visitor Studies* 14 (2), pp. 158–161.
- Norris, P. (2001). *Digital Divide: Civic Engagement, Information Poverty, and the Internet Worldwide*. Cambridge, U.K.: Cambridge University Press.
- Rowe, G. and Frewer, L. J. (2004). ‘Evaluating public participation exercises: A research agenda’. *Science, Technology and Human Values* 29 (4), pp. 512–556.

Spitzberg, B. H. (1983). 'Communication competence as knowledge, skill, and impression'. *Communication Education* 32 (3), pp. 323–329. DOI: [10.1080/03634528309378550](https://doi.org/10.1080/03634528309378550).

Thorne, S. (1997). 'The art (and science) of critiquing qualitative research'. In: *Completing a Qualitative Project: Details and Dialogue*. Ed. by J. M. Morse. Thousand Oaks, U.S.A.: SAGE, pp. 117–132.

Author

Dr Eric Jensen is Associate Professor of Sociology, University of Warwick. He teaches on social research methods, audience research and science communication. Jensen has conducted numerous evaluation studies of science communication, museum, gallery and public engagement activities. He has two forthcoming books with Cambridge University Press, focusing on public engagement events and engagement with wildlife conservation respectively. His research methods textbook published by SAGE is entitled 'Doing Real Research'. Jensen has a Ph.D. in sociology from the University of Cambridge. E-mail: E.Jensen@warwick.ac.uk.

How to cite

Jensen, E. A. (2015). 'Evaluating impact and quality of experience in the 21st century: using technology to narrow the gap between science communication research and practice'. *JCOM* 14 (03), C05.



This article is licensed under the terms of the Creative Commons Attribution - NonCommercial - NoDerivativeWorks 4.0 License.
ISSN 1824 – 2049. Published by SISSA Medialab. <http://jcom.sissa.it/>.