9 Why impact evaluation matters in science communication: Or, advancing the science of science communication

Eric Allen Jensen

Introduction

Science communication activities have different agendas, audiences and venues, but most share the goal of making scientific or technical knowledge and research more accessible for public audiences to understand, discuss or debate. But this leaves open the fundamental questions: What counts as effective science communication? What difference is our science communication making? How can we measure whether the communication approach was effective at developing impact? These questions are fundamental to the science communication enterprise (see, for example, National Science Board [2006]), as their answers provide the pathway to improvement in practice over time.

Impact is the overall net outcomes or results of an activity or intervention (intended or unintended) for individuals or groups. Note that changes or 'impacts' can be in negative or dysfunctional directions (Jensen, 2015b). Impacts could include, for example, development in learning about a specific topic; attitude change; a greater sense of self-efficacy; enhanced curiosity or interest in a subject; and improved skills or confidence.

Despite over two centuries of public science communication practice, there is no consensus on what counts as 'success' for public engagement and informal science learning initiatives.[1] The lack of good evaluation practice across the field is certainly a key contributor to this state of affairs (Jensen, 2015a). Industry standard evaluation conducted across the science communication field comprise a rogue's gallery of errors and poor practice at each stage in the process from research design to instrument design, sampling, analysis and interpretation. This problem of failing to employ established best practice within social research methods to the challenge of evaluating science communication outcomes for audiences extends to other related fields such as museums. When Davies and Heath (2013: 13) reviewed summative evaluation reports produced by numerous UK museums and consultants with the hope of finding 'golden nuggets' of insight, they instead concluded that evaluation 'evidence used to suggest learning or particular forms of learning can appear fragile at best'.

Indeed, low-quality evaluation evidence, as well as the absence of evaluation or evidence-based design of science communication initiatives has been setting up the global enterprise of science communication for failure over many years. Science communication practitioners are rarely trained to be able to distinguish good from bad evaluation methods, and science communication institutions (including funders) are generally uncritical consumers (and producers) of evaluation research. Generally they simply accept results that align with what they wish to believe, without looking too deeply at the methodological rigour underpinning the knowledge claims.

Of course, measuring the impact of science communication on self-efficacy, learning, attitudes and other outcome variables can be challenging. Measuring such impact often requires direct measurement of visitors' thinking or attitudes before and after the science communication activity. However, this direct pre- and post-impact evaluation is rarely implemented within

1 Science communication may be traced back to the Royal Institution Faraday: kingsciencepublic/2018/10/22/triangulating-the-history-of-science-communication-faraday-marcet-and-smart/

science communication practice. Instead, biased feedback survey questions prompting skewed positive responses from audience members comprise the vast majority of evaluation efforts in science communication globally. Poor survey design is routinely used by consultants and practitioners with decades of experience working for top science communication organisations and with funding from pre-eminent scientific institutions and funders. So why don't these institutions and funders demand better? Why don't they apply the same rigorous expectations of scientific research to the communication activities conducted by and on behalf of the same scientific institutions?

Many excuses have been proposed for the widespread lack of methodological quality in science communication evaluation (Jensen, 2015b). However, it is clear that evaluation and social research methods more broadly have not been prioritised in the training of science communicators, despite the centrality of evaluation to good evidence-based practice. Good science communication requires planning that is rooted in the existing knowledge base for science communication, including both theory and research (Dam et al., 2015). It also requires clear objectives from practitioners at the outset in order to establish communication methods that are logically aligned to the aims. Moreover, evaluation results must be planned into the process in order to inform science communication practice. This kind of evidence-based science communication holds real potential for advancing the field over time, if science communication training and education can be enhanced to enable it.

Conducting effective evaluations that accurately measure the intended outcomes and inform practice requires training in relevant aspects of social scientific research methods such as survey design and qualitative data analysis (Krippendorff, 2013). At a more basic level, however, good science communication evaluation requires clear, realistic objectives as a starting point to designing effective measurement tools.

Developing good, evidence-based science communication from the ground up

In order to evaluate a science communication intervention's effectiveness, one first needs to specify the desired outcomes (short and long term). There is a broad range of implicit aims underpinning public engagement with science. However, these aims are rarely made fully transparent to audiences or even to those involved in conducting the activities. This lack of clarity about aims and the logical connections to the science communication activity at hand is widespread across different types of science communication practice (Jensen & Holliman, 2016; Kennedy et al., 2018).

Science communicators should have the end goal in mind, even if that is distant from the initiative itself. Clear definition of aims at a practical level is essential to establish the foundations for effective public communication practice and evaluation (Holliman, 2017b).

The value of taking a systematic approach to defining the nature and level of outcomes for a given public engagement activity includes the following:

- Enables *assessment of success*. Having transparent goals and aims helps to focus the engagement practice itself and to measure the level of its success. This includes checking whether the activity is reaching intended the types of audiences (Jensen et al., 2015).
- *Improves* engagement practice. Use incoming evaluation evidence to continuously improve methods of engaging audiences.
- *Know the impact of the activity:* The activity may be damaging the aims of the public engagement. High-quality evaluation linked to clear, measurable intended outcomes can ensure that the activity remains on track.

An over-emphasis on outputs only (i.e. what you have done, rather than how audiences have benefited) is a common problem in science communication. Science communicators often assume that their outputs necessarily lead to the hoped-for impacts, thereby limiting the scope for improvement over time. Indeed, 'the possibility of negative impacts, are routinely neglected within science communication evaluation' (Jensen, 2014: 2). It is important to have concrete, realistic outcomes specified in advance, which the science communicator then seeks to translate effectively into practice for the benefit of audiences.

This kind of planning information is essential for establishing the basis for accountability and quality in public engagement with science. The other related key factor is ensuring that public engagement initiatives are evaluated regularly for quality of experience, and at least occasionally for impact. Ideally, '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' (Jensen, 2015a: 1).

Of course, even the most 'successful' science communication initiatives based on our definition above could have implications or results that the initial sponsors or science communicators might find distasteful. This openness in outcomes is inherent in communication. Yet, there is clearly a great deal that science communicators can do to improve the quality and value of their activities for its audiences. Explicitly articulating intended outcomes can help to reveal gaps between science engagement practices and the logical steps on the pathway to achieving valuable engagement aims.

Clarifying aims to set up effective science communication evaluation

Limitations in existing science communication evaluation practices are rooted in science engagement practice (Jensen, 2014) and in the aims practitioners set for those practices. The practical question science communicators should be asking on an ongoing basis is: 'How could I improve my science engagement activities?'

To answer this question, clear aims and evaluation are needed, which should feed into practice in a continual manner to establish an evidence-based approach to science communication.

Good science engagement requires upstream planning and clear objectives, and this is even more so for evaluation. Moreover, there should be a commitment to making improvements in the programme or activity based on what the evaluation reveals. It is important to start with the big picture:

- Why are you doing your public engagement event?
- What do you want to achieve with it?
- How will you know if you have been successful?
- Are your goals clear, specific and realistic?

These probing questions can help inform the design of better public engagement activities, while also setting the activity up for good evaluation. Evaluative thinking is oriented towards making improvements, based on good empirical evidence on what is working and why. There are a number of good reasons to evaluate, including:

- To build a better understanding of target publics, (e.g. needs, interests, motivations, language).
- To inform plans and to predict which engagement or learning methods and content will be most effective.
- To know whether the objectives have been achieved (and why or why not).
- To re-design the approach to be more effective in future.

Good impact evaluation is systematic and thorough. It tells one how and why particular aspects of a science communication activity are effective. It does *not* provide a binary 'good'/'bad' or 'successful'/'unsuccessful' result. This is because a 'successful' project can always develop the good aspects of their practice further. Likewise, there will be specific aspects of an 'unsuccessful' project or method that were ineffective (and should be avoided in future projects).

Understanding audiences for science communication

A surprisingly under-developed aspect of science communication evaluation is establishing the nature of the audiences that attend engagement events and activities in order to identify social inclusion gaps and take participant needs into account. To do this effectively, it is important to gather data about participants as they enter the science communication activity. A recent example showing why this is so important comes from the study entitled, 'Preaching to the scientifically converted' by Kennedy et al. (2018). This study addressed the question: 'Are UK science festivals attracting a diverse and broadly representative sample of the public?' It presents findings from evaluation studies conducted in three major UK science festivals. This included pre-visit survey data collected from a science festival in eastern England (n = 592), in southern England (n = 171) and in northern England (n = 1011).

The study showed that in contrast to its aim of widening access to science engagement, most visitors to the science festivals were already highly engaged in cultural and scientific events prior to their science festival attendance. For example, in the northern England science festival, 65% reported already attending other science festivals, events, or activities prior to their visit. In comparison, the 2014 national Public Attitudes to Science (PAS) survey found that 3% of its national UK sample reported attending a science festival.

The study also showed high pre-visit levels of interest in science amongst science festival audiences. In pre-visit responses for both the southern (88%) and eastern (92%) festivals, visitors agreed they were personally interested in science. The study also found that adult attendees at the science festivals were substantially more highly educated than the UK population as a whole, and science festival attendees were more economically advantaged than the general population. This study's audience profiling revealed disparities in access to science engagement, which could reinforce social inequality. Prior to this study, this key information was not available to science festival organisers.

South African evaluation examples

In this section of the chapter, two examples of evaluation from South African science communication practice are presented.

Evaluating the impact of a South African MOOC

This example focuses on a massive open online course (MOOC) led by Prof. Anusuya Chinsamy-Turan and published by the University of Cape Town. This course was on the theme of 'Extinctions'. In order to develop the impact evaluation of this MOOC, the organisers needed to clarify the most important impact objectives.

In this case, those were primarily learning-oriented objectives, key 'take home' points relating to the course theme of extinctions. Once these impact objectives had been clarified, a set of Likert-type items were developed to evaluate the progress towards achieving those outcomes with participants (Table 1). Each of these Likert-type questions asked for a response on a scale from *strongly disagree* to *strongly agree*, with a neutral mid-point and a 'don't know / no opinion' option. Each of these items was repeated before the course, in the middle and after the end.

By repeating these statements at three different points matched to the same individual, the evaluation was able to show the relative progress of different individuals through the course on the defined learning outcomes.

Table 1: Likert-type items developed to evaluate progress towards achieving outcomes

Likert-type item used to evaluate MOOC impact objectives	**Outcome measured**
'Extinctions in the last 100 years are the result of mostly natural processes, not human activity'.	Understanding of key learning point (reversed)
'Human behaviour is negatively affecting ecosystems'.	Understanding of key learning point
'Biodiversity is a valuable resource for humans to use'.	Understanding of key learning point
'The environment is important to me'.	General attitude statement relating to the theme of the course

Likert-type item used to evaluate MOOC impact objectives	**Outcome measured**
'Understanding past extinctions can be important for understanding the effects of the 6th extinction'.	Understanding of key learning point
'I think the 6th mass extinction is already underway'.	Understanding of key learning point
'I think there is little humans can do to prevent the 6th mass extinction from happening'.	General attitude statement relating to the theme of the course (reversed)
'All life on earth will soon come to an end'.	Understanding of nuanced learning point, that is, that the process of extinctions may lead to the end of humans but not to the end of all life (reversed)

World Biotech Tour in South Africa

The World Biotech Tour (WBT) is an ongoing global programme coordinated by the US-based Association of Science and Technology Centers (ASTC), with the goal to demonstrate the relevance, excitement and wonder of biotechnology. It involves students, teachers, science centre professionals and the general public in hands-on activities and discussions about key issues pertaining to biotechnology. In 2017, the WBT travelled to South Africa, with events in different cities across the country. Three science centres from three different cities participated. Sci-Bono Discovery Centre was the lead on this effort, working with Sci-Enza in Pretoria and Cape Town Science Centre.

Different evaluation approaches were used with the different categories of audience for the WBT initiative, with a set of surveys as the primary evaluation approach. The example in this chapter focuses on the Ambassador Programme. For this programme, each of the three science centres assembled a team of high school students (designated as 'ambassadors') to develop and present a biotechnology topic of their choice. A total of 13 students took part in the programme. They were supported by mentors with links to each of the centres, who provided their expertise to help them with their research and presentations. Both ambassadors and their mentors were surveyed as part of the evaluation, taking into account their experience and views. This example addresses the experience and impact for the ambassadors in three of the stages of the programme (pre, mid, and post-programme).

To evaluate the impact and quality of experience for the Ambassador programme during the WBT in South Africa, the following surveys were designed and administered:

- *Pre-programme survey*: This survey included demographic information, interests and motivations relevant to the programme. It also contained outcome measures that were repeated across each survey to allow for evaluation of change over time (i.e. 'impact'). This included measures of biotechnology and general science knowledge and interests, views about scientists and scientific careers, as well as more general transferable skills about confidence and skills they may develop during the programme.
- *Mid-programme survey*: This survey focused on feedback on the experiences of the programme while it was still ongoing to highlight any concerns/issues that should be addressed by the participating science centres.
- *Post-programme survey*: This survey focused on self-report of programme experiences and retrospective assessments to highlight possible areas of improvements for the programme. Additionally, programme impacts were assessed using items that repeated across each survey to show individual-level impacts.

In all cases, initial results were made available to organisers at each of the participating science centres to allow use of the information with pre-programme and intermediate surveys to allow improvements during the programme.

Most evaluations focus only on quality of experience information. In the case of WBT, such feedback questions were included in the survey design in addition to the repeated measures impact items. The post-programme survey items shown below were designed to assess the value the participant placed on the experience of being an 'ambassador' for this programme.

The 'please explain' follow-up questions shown in Figure 1 only appeared when a negative response was submitted by the respondent. In order to go beyond post-visit quality of experience measurement only, repeated measures (pre- and post-test type) were used to evaluate impact by comparing responses before and after the experience (see Figure 2). Each of the questions was repeated exactly at all three data collection points in order to track change over time (i.e. evidence of impact).

Figure 1

	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Not applicable/ no opinion
Overall, the Ambassador's programme was a poor use of my time	○	○	○	○	○	○	○
	Please explain:						
Overall, I found the content of the Ambassador's programme useful	○	○	○	○	○	○	○
	Please explain:						

Figure 2

	Strongly disagree	Disagree	Somewhat disagree	Neutral	Somewhat agree	Agree	Not applicable/ no opinion
Science is irrelevant to my life	○	○	○	○	○	○	○
Biotechnology helps to solve the worlds problems	○	○	○	○	○	○	○
Science is not for me	○	○	○	○	○	○	○
Scientific knowledge is important for my future career	○	○	○	○	○	○	○
Biotechnology is hard to understand	○	○	○	○	○	○	○
If I wanted to, I could be a scientist	○	○	○	○	○	○	○
Science is boring!	○	○	○	○	○	○	○

The ambassadors' *pre-programme* survey results indicate that this participant group understood *biotechnology* as something that can solve problems and generate positive change, and considered it to have potential for helping people address everyday issues.

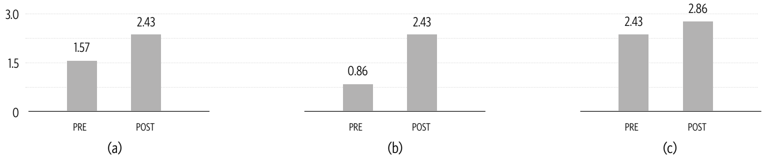
The *mid-programme* results show positive attitudes towards the WBT scheme as participants in this group mentioned advantages of spending time and sharing ideas with scientists and people from other countries. Moreover, the *mid-programme* results evidenced engagement of ambassadors with the WBT programme.

The *post-programme* evaluation shows positive outcomes with the majority of participants highlighting that scientific knowledge is relevant for advancing their future careers. Also, the survey results show that most participants in this group felt they could be scientists if they wanted to, which indicates successful outcomes for the programme in terms of empowering young people. Furthermore, ambassadors in the *post-programme* evaluation indicated that the programme helped them to develop communication and networking skills.

Figure 3a, presents results showing improvement in ambassadors' opinions about their ability to become scientists, and Figure 3b shows positive impact of the programme on their ratings of the importance of science in their careers. In both cases, there is a sharp positive increase from before to after the WBT programme. As a result, the programme has evidently been successful in developing impact on the ambassadors' empowerment and engagement with science. In comparison, the level of impact on attitudes about the relevance of science to ambassadors' daily lives was much less pronounced (Figure 3c). This indicates that the programme is more effective in boosting scientific self-efficacy (the belief in one's capacity to engage with science) than demonstrating to ambassadors the relevance of science to their lives.

This example shows the distinction between widespread quality of experience evaluation, on the one hand, and impact evaluation, on the other.

Figure 3: (a) Self-belief in capacity to be a scientist; (b) Importance of science for future career; (c) attitudes about the relevance of science to daily lives



Evaluating science communication impact

Key challenges in evaluating impact apply to both offline and online science engagement:

- Defining the intended impacts for a particular initiative or activity is topic-, audience-, sponsor- and context-specific.
- Clarifying distinctions between exposure, involvement and impact is essential. Often practitioners unintentionally conflate these goals, undermining any effort at impact measurement (e.g. see Jensen, 2015a).
- There may be a gap between explicit and implicit aims, and motivations for the public engagement with science activity (Jensen & Buckley, 2014; Jensen & Holliman, 2016).
- Impacts may be delayed and unfold over time (Jensen et al., 2017).
- Impacts may emerge due to factors after the initial activity/event/content (i.e. what is sometimes called a 'sleeper effect').
- Impacts can be modulated by the socio-economic profile of public engagement participants (Jensen, 2013).
- Measuring long-term impact can be demanding in terms of both expertise and resources (cf. Jensen & Lister, 2015).

It is clear that continuous evaluation practice tied to real-time results can enable science communication organisations to develop activities that are more likely to deliver positive impacts. In light of the barriers science communication organisations face when working to establish high-quality evaluation, current technology linked to good methods can offer a valuable way forward. Given the practical barriers of required expertise, time and resources, continuous impact evaluation can seem like an insurmountable challenge. Yet recent technological improvements have created new means of gathering and analysing ongoing quantitative and qualitative survey-based evaluation using automation (cf. Jensen, 2015a). While social scientific expertise is always required at some points, an automated system can run in an efficient way to provide insights to science communication organisations on an ongoing basis. One example of such an initiative to establish robust technology-enhanced impact evaluation is ZooWise.[2] The ZooWise initiative provides sector-wide, multilingual and widely usable evaluation tools and metrics for zoos, aquariums, botanical gardens, national parks and other nature-oriented public engagement organisations. A similar initiative is ramping up for science communication, called SciWise.[3] This joins other initiatives such as COVES[4] that are aimed at establishing robust methods for sector-wide evaluation.[5]

Conclusion

Developing more effective evidence-based science communication practice will require greater commitment to robust evaluation and making changes to practice on the basis of such evaluation. To begin this process, dramatic improvement in survey design across the international field of science communication is needed (Jensen, 2014). Once good evaluation instruments are established, accurate methods of gathering and analysing data are needed. Throughout this process, it is important to keep in mind that 'success' should not be assumed. 'Given the complexity of science communication interactions – bringing together multiple

2 www.zoowise.org

3 www.sciwise.org

4 www.understandingvisitors.org

5 Practical examples and 'top tips' on evaluation can be accessed at: practicalevaluation.tips

individuals' values, assumptions, world views and meaning-making processes – the remarkable scenario is when positive outcomes are achieved' (Jensen, 2015b: 1). This means that science communication evaluation efforts should start from a neutral standpoint, open to the possibility of both positive and negative outcomes. This standpoint makes it most likely that the evaluation will be useful in highlighting where improvements are needed to make a science communication activity more successful.

The systemic failures in science communication practice must be brought into the light through robust evaluation in order to reveal the pathway to better practices and impacts. At the same time, positive impacts developing from effective practices must be identified systematically in order to develop even more beneficial outcomes. High-quality impact evaluation can be combined with theoretically and empirically informed planning process and ongoing critical self-reflection to enable evidence-based science communication to achieve new heights of positive impact for society (Holliman, 2017a).

References

Dam, F., Bakker, L., Dijkstra, A. & Jensen, E. A. (in press). *Science Communication: A knowledge base*. World Scientific Publishing.

Davies, M. & Heath, C. (2013). *Evaluating Evaluation: Increasing the impact of summative evaluation in museums and galleries*. http://visitors.org.uk/evaluating-evaluation-%E2%80%90-increasing-impact-summative-evaluation-2013/

Davies, M. & Heath, C. (2014) 'Good' organisational reasons for 'ineffectual' research: Evaluating summative evaluation of museums and galleries, *Cultural Trends, 23(*1), 57–69. doi: 10.1080/09548963.2014.862002.

Jensen, E. (2013). Re-considering 'The Love of Art': Evaluating the potential of art museum outreach. *Visitor Studies*, 16(2), 144–159. doi: 10.1080/10645578.2013.827010.

Jensen, E. (2014). The problems with science communication evaluation. *Journal of Science Communication,* 1(2014), C04. http://jcom.sissa.it/archive/13/01/JCOM_1301_2014_C04/JCOM_1301_2014_C04.pdf.

Jensen, E. (2015a). Evaluating impact and quality of experience in the 21st century: Using technology to narrow the gap between science communication research and practice. *JCOM: Journal of Science Communication*, 14(3), C05. https://jcom.sissa.it/archive/14/03/JCOM_1403_2015_C01/JCOM_1403_2015_C05

Jensen, E. (2015b). Highlighting the value of impact evaluation: Enhancing informal science learning and public engagement theory and practice. *JCOM: Journal of Science Communication*, 14(3), Y05. https://jcom.sissa.it/archive/14/03/JCOM_1403_2015_Y05.

Jensen, E. & Buckley, N. (2014). Why people attend science festivals: Interests, motivations and self-reported benefits of public engagement with research. *Public Understanding of Science,* 23(5), 557–573. doi: 10.1177/0963662512458624.

Jensen, E. A. & Lister, T. P. (2015). Evaluating indicator-based methods of 'measuring long-term impacts of a science center on its community'. *Journal of Research in Science Teaching*, 53(1), 60–64. doi: 10.1002/tea.21297.

Jensen, E. & Holliman, R. (2016). Norms and values in UK science engagement practice. *International Journal of Science Education – Part B: Communication and Public Engagement,* 6(1), 68–88. doi: 10.1080/21548455.2014.995743.

Jensen, E. A., Kennedy, E. B. & Verbeke, M. (2015). Outreach: Science festivals preach to the choir. *Nature,* 528(193). [DOI: 10.1038/528193e]

Jensen, E., Moss, A. & Gusset, M. (2017). Quantifying long-term impact of zoo and aquarium visits on biodiversity-related learning outcomes. *Zoo Biology,* 36(4), 294–297. doi: 10.1002/zoo.21372.

Kennedy, E. B., Jensen, E. A. & Verbeke, M. (2018). Preaching to the scientifically converted: Evaluating inclusivity in science festival audiences. *International Journal of Science Education Part B: Communication & Engagement,* 8(1), 14–21. doi: 10.1080/21548455.2017.1371356.

Krippendorff, K. (2013). *Content Analysis: An introduction to its methodology* (3rd edn). Pennsylvania: Sage.

Holliman, R. (2017a). Supporting excellence in engaged research. *Journal of Science Communication*, 16(05), C04.

Holliman, R. (2017b). Assessing excellence in research impact. National Coordinating Centre for Public Engagement Blog. Bristol: NCCPE.

Holliman, R. (2017c). Supporting excellence in engaged research. *Journal of Science Communication,* 16(5). doi 10.22323/2.16050304.

National Science Board (2006). Science and engineering indicators. Washington, DC: US Government Printing Office.