Evaluating the long-term impact of live science demonstrations in an interactive science show

Wendy Sadler
Cardiff University (UK)
sadlerwj@cardiff.ac.uk

Abstract

Science shows are a popular format for communicating science that are used widely across the world, yet there is little literature about the long-term impact they may have. This research investigated the short- and long-term impact of a specific science show called 'Music to your ears', which was written and performed throughout the UK for students aged 11-16 years on behalf of the Institute of Physics. The impact is measured using the immediate reaction to the show, the number (and type) of demonstrations that were recalled long-term, and the application of any memories since seeing the show.

Quantitative and qualitative data were gathered using questionnaires immediately after the show and from focus groups that were held two and a half years later. In addition, interviews with professional presenters were conducted to allow for triangulation of data. Five science demonstration categories were developed to describe the essence of a demonstration; curiosity (C), human (H), analogy (A), mechanics (M) and phenomena (P). These categories were then used as a framework to discover which had the biggest impact over a long period of time.

It was found that even after two and a half years, almost 25% of demonstrations from the show could be recalled without prompting. When prompted with verbal and visual clues, the group tested could recall more than 50% of the demonstrations from the original show.. In addition, around 9% of the demonstrations were recalled in an alternative context to the show, suggesting that some cognitive processing may be happening with the most memorable elements of the show.

1. Introduction

The demonstration lecture has been an informal method of science communication for many years [1]. Michael Faraday presented science lectures for the public from the 1820s at the Royal Institution and firmly believed they were an important way to engage wider audiences and children with science [2]. Faraday recognised that using live demonstrations was an essential part of communicating science. Moving on to more recent years, there has been a steady growth in the number of science centres and other performers operating worldwide, many of whom use science shows as part of their mission to inspire and engage audiences. Despite this, there is a distinct lack of published literature about their long-term impact or detailed analyses of the demonstration types that are most effective. This study aims to use the triangulation of data from audiences and professionals in an attempt to learn more about the potential impact of demonstrations within the science show format.

2. Methodology

2.1 Research tools and suitability for this study

The two main paradigms of social science research are positivistic and phenomenological [3]. Positivistic research tends to be objective and deals in mainly quantitative methods that assume people behave in a way that can be reproduced to obtain the same results over and over again. Phenomenological research on the other hand accepts that people are affected in some way by being involved in the research and that the researchers themselves cannot separate their own views and beliefs when conducting and even designing the research. This study contains elements of both with an acknowledgement of the limitations given by my close connection to the show being studied. The three stages and purposes of each research tool used are detailed in Figure 1.

Figure 1: Chosen research methods and purpose

Research method	Sample size	Type of data	Purpose of collecting data
Initial questionnaires	171	Quantitative and qualitative	To assess short-term impact on a large sample size To shape hypotheses for the focus group discussions
Focus Groups	8	Qualitative	To assess long term impact To test out hypotheses of memorable demos
Interviews with professionals	6	Qualitative	To compare my emerging hypothesis of demo categories with other professionals to see how generalisable the categories might be

2.2 Questionnaires and demo categories

The questionnaires consisted of both open and closed questions and were intended to give a large sample of responses from the audience immediately after seeing the presentation. The sample size was fairly large (n=171) and some of the data collected was numeric and suitable for basic statistical analysis. The questionnaires also attempted to use the grounded theory method of research where there was no initial hypothesis in mind and I hoped that the responses from the open-ended questions would help lead my research and define the direction of the focus groups. Responses from these questionnaires formed the demo category list shown in Figure 2.

Figure 2: Demonstration Categories by characteristic

Category code	Category name	Definition	The audience…?
C	Curiosity	Something weird, a piece of equipment never seen before, something counter-intuitive, a challenge to the audience	Are surprised
H	Human	A volunteer is used, all of the audience take part in an experiment, something personal is learned, something funny happens to volunteer	Interacts
A	Analogy	A visual representation of something that is usually invisible, using body language to draw a mental picture, using models	Understands
M	Mechanics	How things work, taking things apart and seeing what is inside, how the science is applied in real life things, how to make simple things, in general applied science – technology	Contextualises
P	Phenomena	A chance to see a scientific phenomena happening live, may use equipment not readily available, not necessarily weird, illustrates the basic science using a demonstration rather than words	Experiences

Many demos have dual (or even triple) purpose but there is usually one primary reason for putting it in the show.

2.3 Focus groups
The focus groups were designed to follow up the trends that came out of the questionnaire analysis, and therefore they were not scripted until analysis of the questionnaire data was complete. Due to the long period of time that had passed the focus groups were fairly small in size (n=8) but interestingly consisted of a range of students – some who had gone on to study A level physics and others who had not. This allowed for an unintended variable to be examined in terms of those who had chosen to go on and study physics. The text from the focus groups was transcribed in full and then content analysis was conducted to look for the language used when describing demonstrations from the show. The frequency and type of words used was coded so it could be numerically analysed.

2.4 Interviews with professionals
The British Interactive Group Event is an annual conference for anyone working in the field of interactive communication. I organized some semi-structured interviews and focus groups with a total of six professional presenters at this event which were transcribed. The aim of this extra data collection was to try and establish how my process of demo categorisation fitted within other professional opinions, and also to enrich the data from my small study with some wider comments about the field of science shows in general.

2.5 Why use triangulation of methods?
The use of triangulation can overcome the potential bias and sterility of a single method approach [3]. I felt that this would be a useful tool to try and counteract the closeness of my own expertise to the focus of the research. In addition, triangulation can lead to a greater validity and reliability [4] which I felt was important if I wanted to provide general guidelines for best practice based on the results of the project. Specifically I used a combination of 'method triangulation' and 'data triangulation'. Data triangulation is where data is collected at different times or from different sources in the study of a phenomenon. Method triangulation is where both quantitative and qualitative methods of data collection are used [5].

3. Summary of Results

3.1 Overview of results
Content analysis of the combined data gives the following results in terms of how frequently different categories of demonstration are mentioned in each case.

Figure 3 : Triangulation of data and the prioritising of demo categories

Questionnaires (Initial impact)	Focus Group (long-term memory)	Professional presenters (presenter perspective)
1. Curiosity (33%)	1. Curiosity (25%)	1. Curiosity (50%)
2. Human (25%)	2. Mechanics (18%)	2. Human (25%)
=3. Analogy (17%)	3. Human (14%)	=3. Analogy (12.5%)
=3. Mechanics (17%)	4. Phenomena (11%)	=3. Phenomena (12.5%)
5. Phenomena (8%)	5. Analogy (6%)	5. Mechanics (0%)

Clearly all the data agrees that 'Curiosity' based demos have the most immediate impact, are most memorable and are most important to presenters. In addition the 'Human' category comes in second on two occasions suggesting that this is the next most important type of demo to both audiences and presenters. The only anomaly where the presenters do not mention a category that is fairly popular both in the short and long-term is 'mechanics'. This was not mentioned at all in the presenter interviews (demos about 'how things work') and yet scores reasonably well in the other data sets.

3.2 Discussion
Without any verbal or visual prompts, members of the focus group managed to recall around 25% of the demonstrations used in the show after a period of two and a half years had elapsed. With some visual prompts the groups managed to recall over 50% of the demos used. This was a higher than expected amount of recall from the focus groups after two and a half years had passed but supports

the suggestion that people can recall demonstrations some time after a show is over [6]. There is evidence that around 9% of the memories suggested related links being made to things they saw in the show. There is sometimes criticism that events like science shows have a short-hit lifetime that is quickly forgotten, but this data suggests that with some audience members at least, this is not the case.

4. Applying the research to best practice

Based on this research it is recommended that science show professionals ensure a mixture of the **CHAMP** demo categories within their presentations as there is evidence to suggest different types of audience respond to different categories of demonstration. However, some generalisations can be made:

- **Curiosity** type demos seem to be universally popular regardless of the audience and have a high impact rate for short and long-term recall
- **Human** angle demos are also highly memorable (though not quite as much as 'curiosity')
- **Mechanics** type demos are more popular with audiences that are already interested in science. Mechanics demos are also the type that are most likely to help people relate the show to other contexts
- **Analogy and Phenomena** demos are useful educational tools, but they tend to have less short and long-term impact.

In addition, the data suggests that short-term impact is likely to be similar to that which is remembered in the long-term. This means that when there is a lack of resources for longitudinal studies, it may be possible to extrapolate form the short-term impact to make a hypothesis about the kind of things that are likely to be remembered over a longer period of time.

References

[1] O'Brien, T. (1991). The science and art of science demonstrations. Journal of Chemical Education, 68, 933.

[2] James, F. A. J. L. (2002). 'Never talk about science, show it to them': The lecture theatre of the royal institution. Interdisciplinary Science Reviews, 27(3), 225-229.

[3] Hussey, J., & Hussey, R. (1997). Business research. London: Macmillan Press Ltd.

[4] Denzin, N. K. (1970). The research act: A theoretical introduction to sociological methods. Chicago: Aldine.

[5] Easterby-Smith, M., Thorpe, R., & Lowe, A. (1991). Management research: An introduction. London: Sage.

[6] Burns, T. (2003). Science shows: Evaluating and maximising their effectiveness for science communication University of Newcastle, Australia. (unpublished PhD thesis)