Tutorial: Public Engagement Through Audio Internet Experiments

Trevor J. Cox
Acoustics Research Centre, University of Salford, Salford M5 4WT, UK, t.j.cox@salford.ac.uk

This tutorial paper details experiences of four public engagement projects that have communicated acoustic science to lay audiences using web experiments. Recent developments in personal computers, the Internet and software platforms offers new and exciting opportunities for engaging publics because technologies routinely allow the reproduction of sound. The projects are psychoacoustic experiments run via the Internet (there are an increasing number of psychology experiments mediated via the web).

I. INTRODUCTION

Public engagement is a broad term used to describe a myriad of activities involving experts and publics, which are a two way processes involving interaction and listening, with the goal of generating *"awareness, enjoyment, interest, opinion-forming, and understanding"* [1]. Public engagement in science has been increasing in recent decades. The engagement is sometimes driven by a desire to reverse declining numbers of people wanting to pursue scientific qualifications and careers. Another common motivation for engagement is to ensure that the public can make informed choices about scientific developments, especially controversial issues such as vaccination, nanotechnology and GM crops.

Frequently, engaging the public is done face-to-face, but now the widespread availability of digital technologies such as multimedia computers, laptops, smart phones and tablet computers enables new forms of engagement. There are over a billion computers in use across the World[2] and by 2012 it is predicted there will be another 300m smartphones in the world[3]. Most of these devices have the capability of reproducing sound and many have the capability of recording sounds. Consequently there is an opportunity to engage the public around acoustic science and engineering through these technologies. This paper intends to detail lessons learnt from undertaking four engagement projects.

In the web experiments people audition and make judgements about recorded sounds. Since the 1990s, the use of psychological experiments mediated through the Internet has been increasing[4]. Web experimentation using sound was initially uncommon because old computers were not routinely sold with sound capabilities and the Internet had limited bandwidth. However, technology has improved over recent years and it is now possible to carry out some psychoacoustic tests across the Internet. In this paper, the experience of running four different experiments are discussed and examined. In part to reflect on the methodology, highlighting the advantages and disadvantages inherent in carrying out psychoacoustic experiments across the Internet, but mostly to discuss the possibilities for engaging the public as active participants and to communicate acoustic science.

Over the last decade in the UK, there has been an increasing focus on not just portraying the facts of scientific research but also in getting the public to better appreciate the process of scientific investigation. Evidence for this include: the move within schools to emphasis Science Literacy through How Science Works in the science curriculum since 2006[5]; a bestselling popular science book about scientific methodology[6] and the So You Want To Be A Scientist Competition run by BBC Radio 4[7]. Mass web experiments allow publics to participate and therefore see some aspects of scientific method. One of the web experiments has also been used as part of teaching resources used to support the teaching of How Science Works[8]. Naturally, the web experiments also offer the opportunity for more traditional information giving, raising awareness and understanding of the science behind the experiments, and hopefully being interesting and enjoyable to participate in.

The paper starts by presenting examples of engagement activities. These sections give a sense of the activity, the type of engagement, the reach and lessons learnt. Finally, the paper brings together common learning across the different projects.

II. EXAMPLES

A. Internet experimentation

Four large scale psychoacoustic experiments have been carried out via the Internet: (i) The hunt for worst sound in the world. (ii) The search for the funniest whoopee cushion sound, which was done as part of a fundraiser for a comedy charity. (iii) What makes a scream horrible? which was carried out around Halloween for the Manchester Science Festival. (iv) Can a computer laugh naturalistically? in collaboration with New Scientist magazine to accompany a feature of laughter[9]. The experiments are hosted at www.sound101.org.

The desire to generate mass participation shaped the subjects for the experiment. The experimental designs had to be simple, unambiguous and appealing as these were aimed at the general public. From a user's perspective, the experiments were as follows. When users first went to the website, they were asked for a few details about themselves: their gender, age (within 10-year ranges) and location. This was to give contextual data to enable voting patterns to be better investigated. In future experiments it might be a good idea to ask for some simple description of the sound reproduction system, e.g. laptop loudspeakers, headphones, computer loudspeakers etc., to allow statistical analysis to control for this factor. A cookie was used to temporarily store contextual information on the subject's computer and this information was written to a mySQL database along with the ratings. The IP address of the computer was also stored, and allowed us to analyse how many votes were cast by each user. The experiments were coded in PHP and Flash.

Next participants were presented with the "sound-check" screen to ensure the sound on the computer was turned on and that the reproduction level was reasonable. A sample of speech was presented which said: *"set the volume level so you can hear me speaking clearly, as though I was having a conversation with you"*. This sound-check screen was also intended to help reduce the variation in the volume levels between listeners.

To reduce problems with computer software, the sound files were played using Flash. Flash is a plug-in present on virtually every computer, and this reduces problems such as incompatible media players. As the sounds are broadcast over the Internet, it is vital that the ownership and copyright of any material used is determined. Some sounds were recorded ourselves, some came from libraries of sound effects which allowed broadcast for education purposes and others came from web repositories such as www.freesound.org which allowed re-use under appropriate Creative Commons licences.

Next, participants came to the voting screen, an example of which can be seen in Figure 1; from the screams experiment. Users pressed play, listened to the sound and then voted using the pumpkins at the bottom of the screen. A key to making web experiments popular is to make the interface simple, intuitive and enjoyable to use. Consequently, a direct scaling method on an ordinal scale was used. It would have been possible to use a slider or similar and obtain scale data, however the use of the visually-appealing pumpkins was preferred as being less ambiguous to the user. The use of a short ordinal scale has consequences for the statistical tests used as the voting patterns are distinctly non-normal.

Figure 1. The voting screen used for the scary screams website (text enlarged).

In 2005, the first experiment examined the worst sound in the world. This allowed users to audition as many sounds as they wanted; there were 34 sounds in the database. So the voting screen would keep reappearing with a new sound until the participant got bored. On average people auditioned 11.5 sounds, although the modal number of sounds auditioned was 1. This open-ended approach generated emails from users asking how many sound files they were meant to audition which indicated problems with this approach. Subsequent experiments had users audition a fixed number of sounds (6 and 7 have been used). These sounds were chosen randomly from the database of sounds available. In the scream experiment, 83% of participants rated all 6 sounds and for the laughter experiment 92% rated all 7 sounds. In contrast, for the open-ended worst sound in the world experiment only 68% and 64% of participants rated 6 and 7 sounds respectively. This indicates that setting a limit on the participation, which is shown on the screen to users *"please rate sound 1 of 6"* is probably the best approach.

Consequently, our experience is that most participants will engage with the experiment for a much shorter time than is achieved in laboratory measurements. This limits the experimental methods that can be used; it is not possible to carry out a large scale paired comparison experiment with each participant auditioning many sounds. In contrast, a survey of 35 non-acoustic web experiments found an average participation time of 22 minutes[10], however these experiments had far fewer participants than our web experiments.

After voting was completed, a results page was given. As the primary intention of the experiments was to achieve public engagement, it was vital that the participant gained something from their contributions and were not used as subjects simply to be experimented on. For example, showing how the participants' own responses compared to the average person allowed them to learn a little about their own reactions to sounds. In addition, brief videos were used to communicate some acoustic science behind the sounds and how people respond to them. The feedback has to be chosen carefully so as to not to bias future ratings.

Evaluation of the web sites has been limited to quantification of reach. Unfortunately, funding to run these web experiments was not accompanied by resources to allow an evaluation of the activities. However, the large number of participants who took part in the experiments are indicative that at least the activities were enjoyable.

1. Methodological concerns

If the desire is to generate meaningful scientific data as well as to engage the public in a enjoyable activity, then the methodological concerns surrounding internet experimentation need to be considered. Using the web to run experiments offers a number of advantages over laboratory tests, but also poses a series of methodological challenges[4]. One of the key advantages is that the internet allows access to a much larger and culturally and demographically diverse sample of people than can be easily achieved within a laboratory. Table I shows the total number of people who have taken part in the experiments and Figure 2 gives demographic data for the various experiments. However, there will still be biases in the sample and it may not be representative of the population of interest[11]. It is interesting to note from a public engagement perspective, that many young adults are reached by the experiments, which is a group that is rather infrequently targeted by science communication which more often focuses on school pupils and families. Returning to other documented research advantages of web experiments: participants tend to be more highly motivated than people taking part in laboratory experiments[11]; the absence of the experimenter removes a possible source of bias in the experiment[12], and it is possible to reach rare and specific participant populations[4] although this was not the aim in the sound experiments.

		Number of ratings	Participants	Current number of participants/month (Oct 2010)
Worst sound in the world	Dec 2005	2.2M	189,000	1300
Screams	Oct 2009	63,000	10,500	900
Whoopee cushion sounds	March 2009	220,000	37,000	1100
Laughter	July 2010	63,000	9,000	3000

Table I. Number of participants in each of the web experiments

The disadvantages associated with web experiments have also been extensively discussed in the literature. Drop out is a problem as indicated before, but this can be compensated for by the large sample sizes achievable[12]. There are concerns that the same person may repeat the test, but by monitoring IP addresses or using cookies this effect can be reduced. There are also concerns that non-serious responders will affect the results. However, research has shown that these are not serious problems[11].

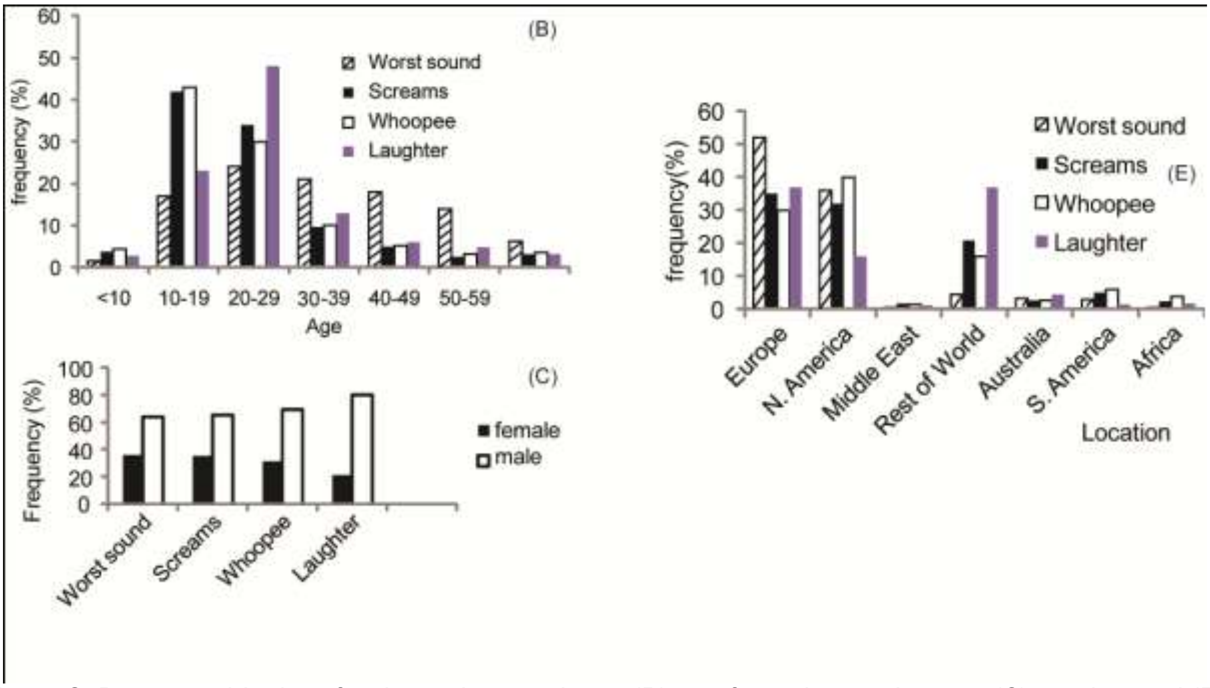

Figure 2. Demographic data for the various projects: (B) age for web experiments; (C) gender, and (E) location for web experiments.

Aligning the experiment to public engagement influences what sort of experiments can be carried out. The need to sell the experiment into conventional media and generate traffic to the website means the idea behind the experiment has to be expressible in a popularist way with a journalistic angle. For instance, the first experiment was sold to the media as the hunt for the worst sound in the world, which naturally meant the website had to be populated with a reasonably large number and range of sounds. However, this meant the experiment was rather unfocussed and made the scientific analysis more difficult. The media were very interested in the idea of finding the worst sound in the world, but in many ways the rank order of sounds produced was rather meaningless. The rank order is dependent on factors such as the quality of recordings and the typical context of listeners. However, on the back of this media-friendly concept it was possible to make more reliable assessment of relative changes, for instance what happens when the website changed colour, and this is where the scientific experimentation becomes more useful and valid.

Environmental factors that may influence web experiments are particularly problematic for sound. There is no calibration of reproduction levels and the frequency and quality of the reproduction equipment varies between participants. In most perceptual experiments, this will cause a significant additional error, but if this is a random error, and the effects are linear, then by getting sufficient numbers of subjects, it is possible to look for underlying trends for the average listening conditions. However, interactions can cause problems. For example, if there is an interaction between listening level and a context variable, say age, then this could also introduce a bias into the results. However, this bias can be minimized by using appropriate analysis. Overall, the lack of calibration and control over sound reproduction is a significant drawback to carrying out acoustic experiments on the web and limits what can be tested.

Furthermore, the context of listening is uncontrolled as participants could be auditioning sounds in many different situations. However, in one sense listening context is no more artificial than is used for many laboratory experiments. To subjects in perceptual experiments, listening rooms and anechoic chambers are often strange, artificial spaces. Indeed some advocates of web experiments argue that a key advantage is allowing participation within familiar environments[4].

For non-acoustic experiments, Musch and Reips[10] surveyed researchers who had carried out 18 experiments where people had compared the results from web and laboratory experiments. They found that *"Almost all experimenters observed complete or good agreement between their web and lab data."*. However, further work is needed to confirm where this is also true for sound experiments.

III. MORE LESSONS LEARNT

B. Design

All projects had to adhere to good application and web site design to be successful. Interfaces needed to be simple to use and based on an understanding of user requirements. In our experience simplicity and ease of use are more important than complex functionality and technical impression.

C. Publicity

Conventional media, newspaper, television and radio were found to be good and reliable ways of publicising the projects. It also affords an opportunity to communicate more acoustic science to a wider public. While word-of-mouth can lead to websites becoming very popular, it is risky to set up a web-based public engagement activity and hope people will find it. Monitoring the traffic to the web experiments showed that most participants were driven to these on the back of news stories. Therefore, for maximum reach, the projects must be about a subject that will appeal to conventional media outlets.

Typical criteria for what makes a good news story helped inform the development of the activities. A story is more likely to interest the media if it: (i) is new to readers and reveals something new about the world; (ii) passes the "so what" test and has direct relevance to readers; (iii) surprises and provokes reaction; (iv) has appeal to non-scientists; (v) entertains and amuses; (vi) is human-centric, and (vii) is topical. Obscure areas of acoustics may be worthy of study, but a journalist's primary task is to entertain. Often, the news stories were dealt with by science journalists, and they need to be reassured that there was a proper scientific purpose behind the work. If the engagement activity is quirky or unusual then not only is it possible to gain media coverage, but it is more likely that word-of-mouth will get more people to the experiment.

However, selling a story of "we are going to do an experiment on…" is difficult because journalists have nothing new to report. Consequently, it has also been found useful to partner with other organisations who can help gain publicity such as the British Science Association or New Scientist. Connecting with events, for instance the scream experiment and Halloween, gives a timeliness hook which is also useful.

The stories were sold using conventional techniques of press releases and phone calls to media contacts. As stories about sound, they were much easier to sell to radio stations, especially when sound files were made available to use on-air. It helped to have a good picture to sell to newspapers. Because these engagement activities were on-going for a period of time, it was possible for the launch publicity to stretch over many days. In contrast, the story about the results from the web experiments were very short lived. We found it sell these engagement activities as a news item rather than as a feature item in a magazine program.

The projects were targeted at a UK audience because of the funding stream which paid for the work and were directly sold to national media outlets and press agencies. However, interest in the projects stretched further afield. With global online coverage generated via websites, such as BBC online, UK and international newspapers as well as national and international radio. For the whoopee cushion experiment, a public relations agency which specialised in education and science was employed to sell the project

The blogging community has played an important role in publicising the more recent projects through their extensive subscriber lists. With the addition of social media outlets, such as Twitter, Facebook and YouTube, links and related articles were seen by more people via re-tweets, Facebook sharing and video embedding.

It is vital that websites are hosted on a reliable hosting service with sufficient bandwidth to cope with sudden increase in users that publicity can generate. The day the results were announced for the worst sound in the world, 500,000 new votes were cast causing the hosting service to believe the website had been hacked, and causing the site to be put onto a probation server.

D. Public attitudes

These projects offer a range of levels of engagement and touched different aspects of science communication in terms of *awareness, enjoyment, interest, opinion-forming, and understanding*. The first and least involved is the passive participant who engages with the project by reading a story in a newspaper, an online source or hearing an interview on a TV or radio show. It is impossible to put a number on this group but a guideline amount can be estimated knowing the audience for the different media outlets. The worst sound in the world experiment generated the most media interest and the reach was estimated to be hundreds of millions of people.

The next level of engagement involves the more active role of commenting on news stories online, taking part in one of the web experiments, getting involved in radio phone-ins to discuss the project themes, joining and actively contributing to the project's social media networks and helping with online development of the project's methodologies. The number of active participants in the web experiments is given in Table 1; the current total is 2.5 million.

E. Collaborations

An unexpected benefit for these projects have been the collaborations with other academics. The relatively high profiles of the projects means they have attracted the attention of other researchers worldwide. The projects have been contacted on many occasions with requests for their sounds and data.

IV. DISCUSSIONS AND CONCLUSIONS

This paper has recounted experiences and lessoned learnt from engaging the public about acoustic science through web experiments. With the right topic for the tests large diverse audiences can be reached. The scientific validity of the results needs further investigation. Like non-acoustic web experiments, some scientists believe the method is valid, while others believe the experiments are little better than a publicity stunt.

V. ACKNOWLEDGEMENTS

Much of the work described in this paper were funded by the UK's Engineering and Physical Sciences Research Council under their Public Engagement programmes. The author would also like to acknowledge the support of Comic Relief, the British Science Association, New Scientist, Museum of Science and Industry in Manchester and the Manchester Science Festival for supporting the web experiments. We would like to acknowledge the support of Jürgen Trouvain, Grégory Beller, Shiva Sundaram and Jérôme Urbain in setting up the laughter experiment; Jon Hargreaves for recording an extraodinary range of whoopee cushion sounds; Sarah Angliss for recording the worst sound in the world, and Louise Heatley and David Dewhurst who authored many of the experiments.

VI. REFERENCES

1. T. W. Burns, D. J. O'Connor and S. M. Stocklmayer, "Science Communication: A Contemporary Definition," Public Understanding of Science **12**, 183-202, (2003).

2. Reuters, "Computers in use pass 1 billion mark: Gartner" http://www.reuters.com/article/idUSL2324525420080623 (accessed 22 November 2010)

3. R. Sambrook, "The world will look different from outside the BBC," The Guardian, media pages pp2 (2010).

4. U-D. Reips, "Standards for Internet-based experimenting," Experimental Psychology **49** 243-256, (2002).

5. A. Shaha, "The S Word: How science works isn't working in British schools," http://www.newscientist.com/blogs/thesword/2009/12/how-science-works-isnt-working.html (accessed 22 November 2010).

6. B. Goldacre, *Bad Science*. Harper Collins 1-288 (2009).

7. BBC Radio 4, "So You Want To Be A Scientist?" http://www.bbc.co.uk/radio4/features/so-you-want-to-be-a-scientist/ (accessed 22 November 2010).

8. Teachers TV, "How Science Works, " http://www.teachers.tv/series/how-science-works (accessed 22 November 2010).

9. K. Douglas, D. Robson D and T. Cox, "Laughter, the best medicine," New Scientist. **207**(2769) 32-36 (2010).

10. J. Musch and U-D Reips, "A Brief History of Web Experimenting," in M.H. Birnbaum (Ed.), *Psychological Experiments on the Internet*. Academic Press 61-66 (2000).

11. S. D. Gosling, S. Vazire, S. Srivastava and O. P. John, "Should we trust web-based studies? A comparative analysis of six preconceptions about internet questionnaires," The American psychologist **59** 93-104 (2004).

12. A. I. Piper, "Conducting social science laboratory experiments on the World Wide Web," Library & Information Science Research **20** 5–21 (1998).