https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

Demonstrating change from a drop-in engagement activity through pre- and post- graffiti walls: Quantitative linguistics and thematic analysis applied to a space soundscape exhibit

Martin O. Archer[1, 2], Natt Day[3], and Sarah Barnes[3]

[1]Space and Atmospheric Physics, Department of Physics, Imperial College London, London, UK
[2]School of Physics and Astronomy, Queen Mary University of London, London, UK
[3]Centre for Public Engagement, Queen Mary University of London, London, UK

Correspondence: Martin O. Archer
(m.archer10@imperial.ac.uk)

Abstract. Impact evaluation in public engagement necessarily requires measuring change, however this is extremely challenging for drop-in activities due to their very nature. We present a soundscape exhibit, where young families experienced the usually inaudible sounds of near-Earth space, which used a novel method of evaluation integrating pre- and post- graffiti walls into the activity. We apply two analysis techniques to the captured before and after data: 1) Quantitative linguistics — Applying Zipf's law (the power law statistics of words) reveals an increased diversity of language concerning space afterwards, highlighting participants engaged with and reflected upon the sounds; 2) Thematic analysis — Finding and grouping patterns in the qualitative data shows altered conceptions of space around aspects of sound, dynamism, emptiness and electricity, areas highly relevant to the underlying space plasma physics of the sonified data. Therefore, we demonstrate that this novel approach to drop-in activity evaluation has the power to capture change from before to after, and thus short-term impact — specifically in this case showing the power of data sonification in innately communicating science. We suggest the method could be adopted by others in their drop-in engagement activities more broadly.

1 Introduction

Drop-in activities — short, interactive, two-way engagements — tend to form a significant fraction of all non-school public engagement, e.g. $31 \pm 3\%$ of all public activities across the UK's South East Physics Network in 2017/2018 were less than 30 min in duration per individual (Galliano, 2018). Such activities though are difficult to effectively evaluate the impact of, since this necessitates a measure of change (King et al., 2015). While pre- and post- surveys may be one of the most robust methods of impact evaluation in general (Jensen, 2014), these are neither appropriate for nor commensurate with drop-in activities since participants are arriving all the time, the engagement duration is so short, and surveys risk affecting participants' experience (Grand and Sardo, 2017). A number of evaluation tools more suitable for drop-in activities have been reported including feedback cards, graffiti walls, rating cards, and snapshot interviews (e.g. Grand and Sardo, 2017; Public Engagement with Research team, 2019). While these methods are particularly useful in process evaluation, assessing the implementation of the

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

activity, under typical usages (post-activity only) they are limited in their ability to routinely demonstrate change from, and thus the impact of, the engagement on participants in general.

This paper presents a novel implementation of graffiti walls integrated into both the start and end of a drop-in activity. This was a soundscape experience surrounding current space science research that used sonified satellite data. We show that this evaluation method (through its design, data collection, and analysis) can indeed capture immediate impact — changed language and conceptions of space in this case. Appendices include details of statistical and qualitative coding techniques employed throughout.

2 Background

A common misconception is that space is a true vacuum completely devoid of matter and thus there is no activity other than that of the celestial bodies, e.g. planets or asteroids. However, the solar system is permeated by tenuous plasmas — gases formed of electrically charged ions and electrons that generate and interact with electromagnetic fields (e.g. Baumjohann and Treumann, 2012). One such example is the solar wind streaming from the Sun, something which only $58 \pm 2\%$ of the UK adult population are aware of (3KQ and Collingwood Environmental Planning, 2015). The solar wind is highly dynamic and as it buffets against Earth's magnetic field generates plasma wave analogues to ordinary sound at ultra-low frequencies (fractions of milliHertz up to 1 Hz) that play key roles within space weather (e.g. Keiling et al., 2016). This contradicts the common belief, perhaps stemming from school science demonstrations such as the bell-jar experiment (see Caleon et al., 2013, for a nuanced discussion) or even popular culture such as the marketing to the movie 'Alien', that there is absolutely no sound in space due to it being "empty".

Sonification — the use of non-speech audio to convey information or perceptualise data (Kramer, 1994) — can be used to convert satellite measurements of these usually inaudible space sounds into audible signals, simply by dramatically speeding up their playback. This has already been leveraged in public engagement projects for both scientific and artistic outputs (Archer et al., 2018; Archer, 2020). Sonification in general has been applied to various scientific datasets (Feder, 2012). Supper (2014) posits that through the public experiencing data in this way it can grip their imagination and produce sublime experiences because of sound's immersive and emotional nature. These arguments, however, are mostly based on reflections from researchers and artists, rather than through the evaluation of participants' own thoughts and feelings. This paper evaluates the short-term impact on participants of experiencing the sounds of space using pre- and post- graffiti walls.

3 Space Soundscape Exhibit

The space soundscape exhibit was held at the free Science Museum in London (United Kingdom) whose informal learning adopts an inclusive, accessible 'science capital' approach that attracts a diverse range of audiences (Science Musuem Group, 2017, 2020). The exhibit formed part of their 'Summer of Space Season', held in celebration of the 50th anniversary of the Apollo moon landings, for which the museum solicited drop-in space-themed activities aimed at young families. It ran between

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

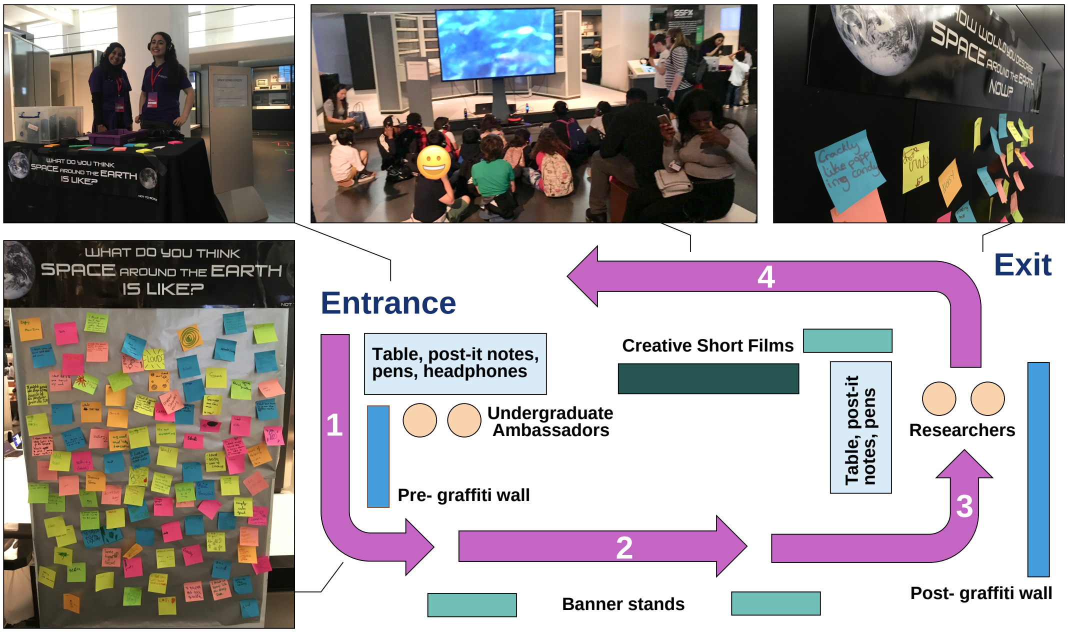

Figure 1. Layout and photos of the soundscape exhibit.

the hours 12:00–16:00 during the May 2019 half-term school holiday over the course of 4 days. Figure 1 shows the layout of the exhibit, which was integrated amongst the museum's usual collections, along with accompanying photos. The activity worked as follows:

1. Museum attendees are invited to participate at the entrance by undergraduate ambassadors. They are first asked to write or draw on a post-it note what they think space around our planet is like. Some younger children required further prompting beyond this broad question however, with ambassadors often asking "what do you think space sounds like?" The participants place their responses on the pre- graffiti wall and are handed bluetooth wireless headphones playing the sounds of space.

2. Participants go on a journey while listening to the sounds, following a set of coloured arrows marked out on the floor. A number of banner stands with further information about the sounds were placed along this path, though it was observed that few people read these. This may be either because participants preferred to listen to the sounds or that it was not clear the stands were part of the experience given the exhibit's location amongst other collections.

3. Near the end of the journey, researchers take participants' headphones and ask them to reflect on what they think about space after having listened to the sounds, again recording their thoughts on post-it notes and placing these on the post-graffiti wall. The researchers would use what they had written or drawn to prompt a short dialogue about aspects of the space environment around Earth and space weather research — a method informed by the 'science capital' research (Archer and DeWitt, 2017).

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

4. Finally, researchers would change the channel on the headphones so that participants could watch on a large TV screen a series of creative short films inspired by and incorporating the sounds featuring epilogue text reinforcing the importance and relevance of space weather research (Archer, 2020). Surprisingly, these artistic films proved much more popular than anticipated.

While graffiti walls are a common evaluation tool, we are unaware of any public engagement activity that has captured data both before and after a drop-in activity using them. This makes the integrated evaluation within the exhibit novel. The space soundscape was experienced by 1,003 people, recorded using a tally counter. No characteristics about the participants were solicited, though the majority were in family groups (approximately three-quarters were children) with some independent adults also. It was observed that in families typically only the children contributed to the graffiti walls and in many cases accompanying adults did not take headphones when offered, perceiving the activity as just for their children. There were 535 and 446 responses (predominantly textual) on the pre- and post- graffiti walls respectively, rates of $53 \pm 2\%$ and $44 \pm 2\%$ — some 3–10 times greater than reported for typical graffiti walls (Public Engagement with Research team, 2019) likely due to their integration into the soundscape activity here.

4 Results and Analysis

The data captured on the pre- and post- graffiti walls are displayed in Figure 2. Two approaches are taken in analysing it, namely quantitative linguistics and thematic analysis.

4.1 Quantitative linguistics

Quantitative linguistics investigates language using statistical methods and has uncovered several linguistic laws that mathematically formulate empirical properties of languages. One of these is Zipf's law — the frequency of words are inversely proportional to their rank, i.e. the distribution is a power law with exponent -1 (Zipf, 1935, 1949). Zipf's law holds well for almost all languages as well as many other human-created systems (Piantadosi, 2014). The Zipf exponent is a measure of the diversity of words and Baixeries et al. (2013) showed that children's exponents become less-negative / shallower with age, demonstrating increasing variety of language and thus linguistic complexity. However, we are not aware of Zipf's law being exploited in public engagement evaluation before.

Figure 3 shows rank-frequency plots of the textual responses to the soundscape before and after the experience. It is clear from these plots that the distributions follow broken power laws (apart from the top word which is of similar frequency before and after), with the break points and exponents being ascertained by a piecewise regression (see Appendix A). Interestingly, the breaks in the two datasets occur at similar ranks namely $\sim$2–3 and $\sim$9–10. While the exponents in the lowest ranked segments are consistent with one another, those in the higher rank segment show clear differences — the after dataset exhibits a much shallower exponent. This indicates significant increased diversity of words resulted following the soundscape, signifying that participants engaged with and reflected on the experience rather than perhaps drawing from common associations concerning

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

Figure 2. Wordclouds (a,b) and drawn images (c,d) from both before (a,c) and after (b,d) experiencing the soundscape.

space. We have therefore demonstrated language change in participants resulting from a public engagement activity through the novel usage of Zipf's law applied to graffiti wall responses.

4.2 Thematic analysis

Thematic analysis (Braun and Clarke, 2006) was used to find, group, and analyse the meaning behind both textual and drawn responses. Instead of using pre-determined qualitative codes, the analysis drew on grounded theory (Robson, 2011; Silverman, 2010), allowing the themes to emerge from the data as outlined in Appendix B. The main themes and underlying codes determined by the first author were:

– **Sound:** an expression of space being either "silent"/"quiet" or "loud"/"noisy"

– **Emptiness:** relating to "nothing" in the "empty" vacuum or conversely filled with material or activity such as "wind"

– **Dynamism:** whether space is slow ("calm"/"peaceful") or highly dynamic exhibiting busy movement

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

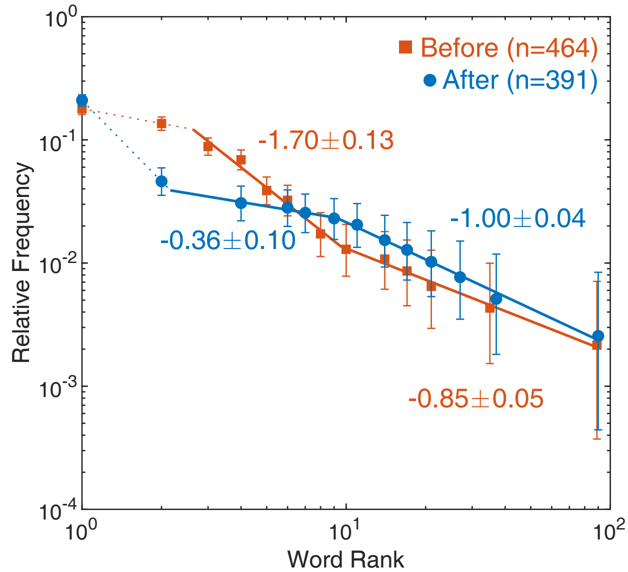

Figure 3. Log-log rank-frequency plot of words before (orange) and after (blue) the soundscape. Power law exponents from a piecewise linear regression are indicated. Uncertainties refer to standard errors.

- **Electricity:** expressions of electrical phenomena
- **Space objects:** references to commonly known celestial bodies (planets, stars, meteors etc.) or artificial spacecraft

An internal consistency test across the five themes gives a Cronbach's alpha of 0.23, indicating they are largely independent of one another as desired. We therefore quantify the number of responses in each theme and qualitative code (cf. Sandelowski, 2001; Sandelowski et al., 2009; Maxwell, 2010) to investigate any changes from before to after the soundscape experience as shown in Figure 4 relative to the total responses (panel a) and within each theme (panel b).

The theme of sound is highly relevant to the activity and was commonly expressed both before and after the activity. Responses before mostly considered space to be quiet/silent ($61 \pm 3\%$ within the theme) but a non-negligible fraction thought it to be loud, which may be due to participants second-guessing the question because of the nature of the activity and/or the phrasing by undergraduate ambassadors. Nonetheless, the overwhelming majority ($97 \pm 1\%$ within the theme) after the experience expressed space to be a noisy environment — a considerable change to beforehand. We note that the theme of dynamism exhibits quantitatively similar results to that of sound — a clear majority ($59 \pm 3\%$ within the theme) thought space to be slow beforehand, whereas the vast majority ($96 \pm 1\%$) consider it highly dynamic afterwards.

The theme of emptiness (including both of its underlying codes) was quite common in responses beforehand, however it was expressed much less often following the soundscape. As a proportion of all responses, space being full was communicated a similar number of times both before and after. In contrast, the prevailing opinion beforehand was that space is empty and this dramatically reduced following the soundscape, both relative to the total responses (from $47 \pm 2\%$ to $2 \pm 1\%$) and within the theme (from $70 \pm 3\%$ to $8 \pm 4\%$).

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

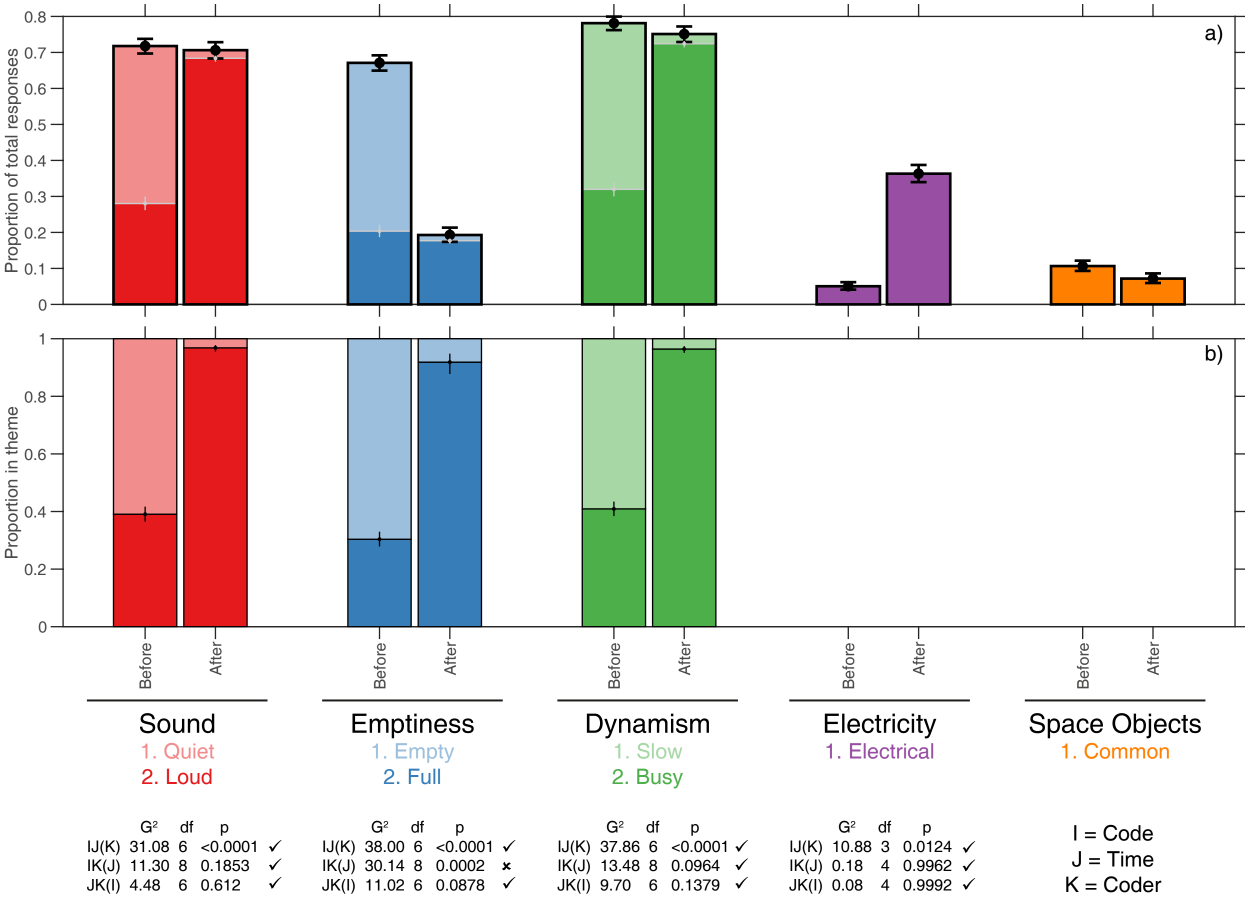

Figure 4. Comparison of qualitative themes and codes before ($n = 535$) and after ($n = 446$) the soundscape experience normalised by total responses (a) and totals within each theme (b). Error bars depict the standard error in proportions. Log-linear analysis statistics of the agreement between coders are also shown for each theme.

There was a clear increase in the proportion of responses relating to electricity following the event, from $5 \pm 1\%$ to $36 \pm 2\%$. While common space objects (typically expressed through drawings) may appear at first glance of Figure 2 to be more frequent before the soundscape than after, as a fraction of the total number of responses this difference is small and not strictly statistically significant ($p = 0.057$).

We checked the reliability of all these trends resulting from the qualitative coding by applying log-linear analysis to a subset of the data additionally coded by the co-authors (see appendices for details). Using the notation that I denotes the qualitative codes, J the time (i.e. before or after), and K the different coders, for the results to be consistent one would expect that the $IJ(K)$ test be statistically significant, constituting the reported trends in codes with time, but the $IK(J)$ and $JK(I)$ interactions should not be, indicating independence from individual coders. These statistics are displayed in Figure 4 for each theme (apart from space objects which was less common) indicating the expected behaviour apart from in the case

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

of emptiness — this theme showed some inconsistency between coders for the "full" code, whereas when only "empty" was considered coders were in agreement ($G^2 = 32.2, 3.42, 2.06$ respectively). Therefore, the main results of the paper are robust and hence we have demonstrated a change in conceptions of space resulting from a drop-in engagement activity.

5 Conclusions

A challenge within public engagement is evaluating the impact of drop-in activities since this necessitates a measure of change that is appropriate to and commensurate with the engagement (Jensen, 2014; King et al., 2015; Grand and Sardo, 2017). We have presented a novel implementation of a common evaluation tool, graffiti walls (e.g. Public Engagement with Research team, 2019), which were integrated both before and after a soundscape exhibit on space science research using sonified satellite data. The pre- and post- graffiti walls provided data on participants' conceptions of space and, through their integration into the activity itself, had much higher response rates than is typical. The captured data was analysed in two different ways.

We investigated the statistical properties of the words expressed using Zipf's law from quantitative linguistics — that the frequency of words in languages typically follow power laws whose exponents give a measure of the diversity of words, where shallower exponents indicate greater variety. The distributions from the graffiti walls showed that the exponent for the top $\sim$2–10 words became significantly shallower from before to after. This demonstrates increased linguistic complexity concerning participants' thoughts about space after the activity. We are unaware of Zipf's law being used in impact evaluation for public engagement before.

We also investigated themes present in the responses, which again yielded significant and robust changes from before to after. While beforehand participants typically expressed common misconceptions of space being completely empty, silent, and with little activity; after experiencing the space sounds they thought space was a noisy and dynamic environment with electrical phenomena present. It is astounding that simply by listening to the sounds these aspects of the underlying space plasma physics were successfully communicated to participants. This therefore demonstrates the power of sonification for audiences, which had been argued by Supper (2014) based on reflections from researchers and artists, however here we have shown it from evaluating participants' experiences directly. The measured changes in conceptions will have been further reinforced by researchers drawing from participants' own reflections in the subsequent dialogues (cf. Archer and DeWitt, 2017).

Overall, integrating existing evaluation tools suitable for drop-in engagement activities, such as graffiti walls, both before and after a drop-in activity can enable practitioners to demonstrate changes resulting from the engagement and therefore its short-term impact. We suggest that our approach, both in terms of data capture and analysis, could be adopted for a range of different drop-in activities beyond just soundscape exhibits.

Appendix A: Statistical techniques

Statistical uncertainties in proportions are estimated using the Clopper and Pearson (1934) conservative method based on the binomial distribution, where standard (68%) errors are shown throughout.

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

A piecewise linear regression in log-log space was used to minimise the sum of squared error between the data and a model made up of a specified number of line segments whose break points could be varied iteratively. This was performed for an increasing number of segments, each time calculating the degrees-of-freedom-adjusted R^2 which accounts for the number of explanatory variables added to the model:

$$\overline{R}^2 = 1 - \left(1 - R^2\right) \frac{n-1}{n-m-1} \tag{A1}$$

where R^2 is the usual coefficient of determination, n is the number samples, and $m = 2s - 1$ is the total number of explanatory variables in the piecewise linear model with s segments. The final model was selected as the first peak in $\overline{R}^2$ with s. Any segments with only two datapoints are later ignored. The statistical significance of the slopes was determined by ANCOVA with a multiple comparison procedure (Hochberg and Tamhane, 1987) quoting standard errors.

Cronbach's alpha (α_C) is a measure of internal consistency based on average inter-item covariances (Cho, 2016). It can be computed as

$$\alpha_C = \frac{k}{k-1}\left(1 - \frac{\sum_{i=1}^{k} \sigma_i^2}{\sigma_T^2}\right) \tag{A2}$$

where σ_i^2 denotes the variance of the item i out of k and σ_T^2 is the variance of the sum over all items. Cronbach's alpha is typically between 0 and 1, where a value of 1 indicates all items essentially measure the same underlying quantity and thus correlate whereas 0 results from items being independent and uncorrelated.

Finally, log-linear analysis is employed to check the consistency of the changes in coding with time across the different coders. This extension of the χ^2 test of independence to higher dimensions uses a similarly distributed statistic, the deviance, given by

$$G^2 = 2 \sum O_{ijk} \ln \frac{O_{ijk}}{E_{ijk}} \tag{A3}$$

for observed O_{ijk} and expected E_{ijk} frequencies (Agresti, 2007). Here we assess conditionally independent models denoted $IJ(K)$, which tests the two-way IJ interaction with the effects of the IK and JK interactions removed. Computationally this calculates G^2 for each level of K summing the results, with G^2 having $(n_I - 1)(n_J - 1)n_K$ degrees of freedom.

Appendix B: Qualitative coding

The qualitative coding process of thematic analysis drawn from grounded theory involved the following steps:

1. Familiarisation: Responses (Figure 2) are studied and initial thoughts noted.

2. Induction: Initial codes are generated based on review of the data.

3. Thematic Review: Codes are grouped together into themes and applied to the full data set. Independence of themes is tested.

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

Theme	Codes	Before ($n = 535$)	After ($n = 446$)
	1. Quiet	234	10
Sound	2. Loud	150	305
	Total	384	315
	1. Empty	250	7
Emptiness	2. Full	109	79
	Total	359	86
	1. Slow	247	12
Dynamism	2. Busy	171	323
	Total	418	335
Electricity	1. Electrical	27	162
Space Objects	1. Common	57	32

Table B1. Number of responses in each theme before and after the soundscape.

4. Reliability: Codes are applied to a subset of data by second coders to check reliability of results.

5. Finalisation: Theoretical interpretation and narrative are formulated from final coding.

Table B1 shows the number of coded responses across words and pictures in each theme and its underlying codes both before and after the soundscape experience. Table B2 shows the number of each coding in the top 16 before (58% of responses) and 15 after (49%) words respectively across the three coders, which is used in the log-linear analysis. The codes' association to the raw data can be found in the supplementary material, along with results from second coders applied to a subset of the data.

Data availability. Data supporting the findings are contained within the article and its supplementary material.

Author contributions. MOA conceived the project and its evaluation, performed the analysis, and wrote the paper. ND and SB assisted with the analysis.

Competing interests. The authors declare that they have no conflict of interest.

Acknowledgements. We thank the researchers (Alice Giroul, Christopher Chen, Emma Davies, Jesse Coburn, Joe Eggington, Luca Franci, Oleg Shebanits) and undergraduate ambassadors (Avishan Shahryari, Cheng Yeen Pak, Christopher Comiskey Erazo, Habibah Khanom,

		Coder 1		Coder 2		Coder 3	
		Before	After	Before	After	Before	After
Sound	1 Quiet	8	0	5	0	8	1
	2 Loud	6	11	5	11	4	5
	None	2	4	6	4	4	9
Emptiness	1 Empty	8	0	6	0	9	1
	2 Full	5	3	0	1	6	7
	None	3	12	10	14	1	7
Dynamism	1 Slow	7	0	5	0	6	0
	2 Busy	7	11	4	12	10	11
	None	2	4	7	3	0	4
Electricity	1 Electrical	2	6	2	7	2	6
	None	14	9	14	8	14	9

Table B2. Statistical comparison of the qualitative coding by different coders.

210 Safiya Merali, Yinyi Liu) who helped deliver the exhibit along with all the staff at the Science Museum (including Becky Carlyle, Imogen Small, Sevinc Kisacik). This project has been supported by QMUL Centre for Public Engagement Large 2016 and Small 2019 Awards, an EGU Public Engagement Grant 2017, and STFC Public Engagement Spark Award ST/R001456/1. M.O. Archer holds a UKRI (STFC / EPSRC) Stephen Hawking Fellowship EP/T01735X/1.

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

References

3KQ and Collingwood Environmental Planning: Space weather public dialogue, Tech. rep., Sciencewise, Science and Technology Facilities Council, RAL Space, Natural Environment Research Council, National Grid, Lloyd's of London, https://www.ralspace.stfc.ac.uk/Pages/SWPDFinalReportWEB.pdf, 2015.

Agresti, A.: An Introduction to Categorical Data Analysis (Second Edition), Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, New Jersey, United States, https://doi.org/10.1002/0470114754, 2007.

Archer, L. and DeWitt, J.: Understanding Young People's Science Aspirations: How students form ideas about 'becoming a scientist', Routledge, London, UK, https://doi.org/10.4324/9781315761077, 2017.

Archer, M. O.: Space Sound Effects Short Film Festival:using the film festival model to inspire creative art-science and reach new audiences, Geosci. Commun., 3, 147–166, https://doi.org/10.5194/gc-3-147-2020, 2020.

Archer, M. O., Hartinger, M. D., Redmon, R., Angelopoulos, V., Walsh, B. M., and Eltham Hill School Year 12 Physics students: First results from sonification and exploratory citizen science of magnetospheric ULF waves: Long-lasting decreasing-frequency poloidal field line resonances following geomagnetic storms, Space Weather, 16, 1753–1769, https://doi.org/10.1029/2018SW001988, 2018.

Baixeries, J., Elvevåg, B., and Ferrer-i-Cancho, R.: The Evolution of the Exponent of Zipf's Law in Language Ontogeny, PLoS ONE, 8, e53 227, https://doi.org/10.1371/journal.pone.0053227, 2013.

Baumjohann, W. and Treumann, R.: Imperial College Press, London, UK, https://doi.org/10.1142/P850, 2012.

Braun, V. and Clarke, V.: Using thematic analysis in psychology, Qualitative Research in Psychology, 3, 77–101, https://doi.org/10.1191/1478088706qp063oa, 2006.

Caleon, I., Subramaniam, R., and Regaya, M. H. P.: Revisiting the bell-jar demonstration, Physics Education, 48, 247–251, https://doi.org/10.1088/0031-9120/48/2/247, 2013.

Cho, E.: Making reliability reliable: A systematic approach to reliability coefficients, Organizational Research Methods, 19, 651–682, https://doi.org/10.1177/1094428116656239, 2016.

Clopper, C. and Pearson, E. S.: The use of confidence or fiducial limits illustrated in the case of the binomial, Biometrika, 26, 404–413, https://doi.org/10.1093/biomet/26.4.404, 1934.

Feder, T.: Shhhh. Listen to the data, Physics World, 65, 20–22, https://doi.org/10.1063/PT.3.1550, 2012.

Galliano, D.: SEPnet Outreach & Public Engagement 2017/18 Reporting, Tech. rep., South East Physics Network, unpublished internal document, 2018.

Grand, A. and Sardo, A. M.: What Works in the Field? Evaluating Informal Science Events, Front. Commun., 2, https://doi.org/10.3389/fcomm.2017.00022, 2017.

Hochberg, Y. and Tamhane, A. C.: Multiple Comparison Procedures, Wiley Series in Probability and Statistics, John Wiley & Sons, Inc., Hoboken, New Jersey, United States, https://doi.org/10.1002/9780470316672, 1987.

Jensen, E.: The problems with science communication evaluation, J. Sci. Commun, 13, https://doi.org/10.22323/2.13010304, 2014.

Keiling, A., Lee, D.-H., and Nakariakov, V., eds.: Low-Frequency Waves in Space Plasmas, Geophysical Monograph Series, American Geophysical Union, https://doi.org/10.1002/9781119055006, 2016.

King, H., Steiner, K., Hobson, M., Robinson, A., and Clipson, H.: Highlighting the value of evidence-based evaluation:pushing back on demands for 'impact', J. Sci. Commun, 14, https://doi.org/10.22323/2.14020202, 2015.

https://doi.org/10.5194/gc-2020-41
Preprint. Discussion started: 13 October 2020

Kramer, G.: An Introduction to Auditory Display, Auditory Display: Sonification, Audification, and Auditory Interfaces, Addison-Wesley, Reading, MA, 1994.

Maxwell, J. A.: Using numbers in qualitative research, Qualitative Inquiry, 16, 475–482, https://doi.org/10.1177/1077800410364740, 2010.

Piantadosi, S. T.: Zipf's word frequency law in natural language: A critical review and future directions, Psychon. Bull. Rev., 21, 1112–1130, https://doi.org/10.3758/s13423-014-0585-6, 2014.

Public Engagement with Research team: Little Book of Evaluation Tools: Curiosity Carnival, Tech. rep., University of Oxford, https://www.mpls.ox.ac.uk/public-engagement/latest/little-book-of-evaluation-tools-curiosity-carnival, accessed: Aug 2020, 2019.

Robson, C.: Real World Research, John Wiley and Sons Ltd., Hoboken, New Jersey, USA, 2011.

Sandelowski, M.: Real qualitative researchers do not count: The use of numbers in qualitative research, Res. Nurs. Health, 24, 230–240, https://doi.org/10.1002/nur.1025, 2001.

Sandelowski, M., Voils, C. I., and Knafl, G.: On quantizing, J Mix Methods Res., 3, 208–222, https://doi.org/10.1177/1558689809334210, 2009.

Science Musuem Group: Inspiring futures: Strategic priorities 2017-2030, Tech. rep., Science Musuem Group, London, UK, https://learning.sciencemuseumgroup.org.uk/wp-content/uploads/2020/04/Inspiring-Futures-Strategic-Priorities-2017-2030.pdf, accessed: Aug 2020, 2017.

Science Musuem Group: Engaging all audiences with science: Science capital and informal science learning, Tech. rep., Science Musuem Group, London, UK, https://learning.sciencemuseumgroup.org.uk/blog/engaging-all-audiences-with-science-science-capital-and-informal-science-learning/, accessed: Aug 2020, 2020.

Silverman, D.: Doing Qualitative Research: A Practical Handbook,, Sage Publications Ltd., Thousand Oaks, California, USA, 2010.

Supper, A.: Sublime frequencies: the construction of sublime listening experiences in the sonification of scientific data, Social Studies of Science, 44, 34–58, https://doi.org/10.1177/0306312713496875, 2014.

Zipf, G. K.: The psycho-biology of language, Houghton Mifflin, Boston, MA, USA, 1935.

Zipf, G. K.: Human behavior and the principle of least effort, Addison-Wesley Press, Boston, MA, USA, 1949.