EVALUATION TOOLKIT

europlanet

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654208.

1 – TOOLKIT OVERVIEW

Europlanet has developed a dedicated evaluation toolkit to empower outreach providers and educators in measuring and appraising the impact of their activities. This toolkit is intended to provide advice and resources that can be simply and easily integrated into normal outreach and education activities.
We are not saying you have to evaluate *all* your outreach and education work! But if you do decide to do some evaluation we want to help ensure it goes as smoothly as possible.
This Section contains brief summaries of all of the subsequent contents within the Toolkit, to help you determine which aspects are of most relevance to you. We'd advise beginning by reviewing the Introduction and the Steps to choosing the right tools sections first. These will give you some general background to evaluation (Why do it? What's the point?) and lead you through key advice in choosing the best tool for the job.
We realise that for most users, evaluation (or even public engagement) is not a major part of your role. If you have any suggestions for how this toolkit can be more useful for you, please get in touch with Anita Heward, Europlanet Outreach Coordinator, at anita.heward@europlanet-eu.org.

TABLE OF CONTENTS

SECTION OVERVIEWS

1.1 — INTRODUCTION

If you're new to evaluation, or want to know more about this specific Toolkit, then start here. We begin with a brief introduction to evaluation generally, as well as the Toolkit specifically. There is also a short description of the intended users for this Toolkit, as well as how it was developed.

We also know it can seem difficult to justify or find the time to do evaluation, so this section includes popular responses to the question '**Why Evaluate?**' and what members of the Europlanet community see as their main purposes when deciding to evaluate an outreach activity.

1.2 — STEPS TO CHOOSING THE RIGHT TOOLS

This crucial advice takes the form of a series of questions which will help you design your approach, and especially make your evaluation as efficient and effective as possible within limited time and resources. We consider each of the following key questions in turn:

- What do I want to find out?
- How will I use that information?
- What resource (time/budget) do I have available?
- How will I analyse the data that I collect?
- Should I consider working in partnership with others?
- Are there any ethical issues I need to consider?
- What methods will I use?

1.3 — TOP TIPS

You're almost ready to get started but before you do, check out our top tips. These tips have been divided into categories according to the various stages of conducting an evaluation (planning, collecting data, coming up with conclusions, report writing and sharing your findings). Following these tips will help you design and conduct a more successful evaluation - one that answers questions you want to know and that helps you communicate and use your findings, so your efforts don't go to waste. We strongly suggest reading them at the start of your evaluation planning!

1.4 — DATA COLLECTION TOOLS

We have prepared an overview table to help with the actual selection of which tool is most appropriate for your situation. This table contains two worksheets:

1) Tools overview - tool name; suitable audience (e.g. primary students, secondary students, etc.); activity type each tool is appropriate for (e.g. interactive workshops, lecture, etc.); description; and whether or not there is a case study example demonstrating that tool in action.
2) Tools detail - the same as above but also includes response type (how people provide their responses - e.g. written, multiple choice, verbal); an indication of the time required for preparation, implementation and analysis respectively; the kind of information that can be gained (e.g. what works, immediate reactions, etc.); and the relevant areas of learning that can be explored using each tool.

	SUITABLE AUDIENCE(S)				APPROPRIATE ACTIVITY TYPE(S)						
TOOL NAME	PRIMARY SCHOOL	SECUNDARY SCHOOL	INTERESTED ADULT	GENERAL PUBLIC	DROP IN (festivals and demonstrations)	INTERACTIVE WORKSHOP	ONGOING SERIES (clubs, courses etc.)	LECTURE PRESENTATION	ONLINE	BRIEF DESCRIPTION	WORKED EXAMPLE INCLUDED?
Tools best suited to use during an event											
OPEN PALM ON CHEST	✓	✓	✓	✓				✓		A simple, low-tech approach to gauge audience responses to multiple-choice questions during a lecture	Image only
PHYSICAL RANKING SCALES	✓	✓	✓	✓		✓				Participants are asked to physically stand along a line representing different levels of experience / attitude etc. (any ranking type question can be asked).	
GRAFFITI WALL	✓	✓	✓	✓	✓	✓		✓		Colour and artistic freedom combine to allow participants to respond to an event in a highly creative way.	Image only
MENTIMETER	✓	✓	✓	✓		✓	✓	✓		A free online interactive presentation tool that allows presenters to quickly (& accurately) gather audience responses.	Yes
Tools best suited to use at the beggining and/or end of an event or activity											
SNAPSHOT INTERVIEWS	✓	✓	✓	✓	✓	✓		✓		Very brief, focused interviews, which are used in conjunction with an event to gather impressions quickly, like a photo of a moment in time.	Yes
PRE/ POST QUIZZES	✓	✓	✓	✓		✓	✓	✓	✓	Brief surveys that are used before & after an event. Ideal for helping you understand whether or not your audiences have learnt key aspects of the content you are trying to convey.	Yes
GEOGRAPHIC LOCATION MAP			✓	✓	✓			✓	✓	Where people come from can be really useful information, however isn't always easy to obtain. This simple and inexpensive technique encourages participants to provide their location data in a fun and visual way instead of a standard survey.	Yes
Tools best suited to use at the end of (or after) an event											
DROPPING, PEBBLES IN BOXES OR STICKERS ON CHARTS	✓	✓	✓	✓	✓	✓	✓			This process is quick, easy, and highly visual, and also provides an opportunity to participate for people who may find reading challenging.	
3 WORDS	✓	✓	✓	✓	✓	✓	✓	✓	✓	A quick, focused way to get a feel for participants' experiences by asking them to describe it in a few short words.	Yes
TARGET EVALUTION	✓	✓	✓	✓	✓	✓	✓	✓		A visual 'bullseye' approach to rating different elements of an event, or other outcomes of interest.	
POST-EVENT SURVEYS	✓	✓	✓	✓		✓	✓	✓	✓	Self-completion questionnaires that are used immediately after an event, workshop or programme.	Yes
PHOTOGRAPHY DIARY	✓	✓	✓	✓		✓	✓			Inviting individuals to select and discuss images that are especially meaningful to them, in their own words, thus providing insights that are otherwise very difficult to attain.	
PEER INTERVIEWS	✓	✓	✓			✓	✓	✓	✓	Peer interviews are a great way to encourage honest opinions, especially from teenagers, through involving participants in interviewing each other, and thereby hopefully allowing a more natural and honest conversation.	
TWEET SENTIMENT VISUALISATION		✓	✓	✓	✓	✓	✓	✓	✓	Automatic online analysis of tweets to help identify participants' reactions to a particular key word or phrase. Must be conducted within a week of the event (whilst the tweets are still "live" on Twitter).	

Based on feedback from members of the Europlanet community the tools are grouped according to when they might best be used (during, beginning/end, or after an event). (We've also included a separate section on Pathways to selecting the most appropriate tools in case the event timing isn't the critical factor in your own activity evaluation). Of course this doesn't mean that some tools couldn't be used at other times (e.g. the 3 words technique is very adaptable, and could really fit into all 3 categories). However we hope that this arrangement will help you to more quickly locate tools likely to be of use to you. The tools are also presented in order of increasing complexity, starting with those that are very simple to implement and moving on to more complex tools (e.g. requiring more preparation or resources such as Wi-Fi).

For some of these tools, members of the Europlanet community have generously provided details of how they applied such tools in their own evaluation of outreach activities. These case study examples include information about the event context, how data was actually collected and analysed and what conclusions were reached, based on the data gathered.

1.5 — ANALYSIS TECHNIQUES

Once you've collected your data, you'll need to analyse it. The following two techniques are some of the most common ways to analyse open-response (qualitative) data. Both of the analysis technique descriptions include worked examples of how to carry out the analyses.

- Word clouds **(*website feedback)** - A fun use of free online tools to easily visualise common themes from data provided by participants in their own words.
- Thematic coding **(*website feedback)** - A process of categorising qualitative data into common themes to assist in identifying patterns within the dataset.

The pre-post quizzes, post-event survey and Thematic coding sections also demonstrate how to conduct basic quantitative analysis. Our scoping discussions with members of the Europlanet community suggested that most users wouldn't be applying detailed **statistical analysis** to any quantitative data collected, however there are plenty of good statistical advice resources available online if you do want to go down this route. One of the most readable authors is Andy Field.

1.6 — RECOMMENDED RESOURCES

In addition to the content provided within this Toolkit there are plenty of other excellent evaluation resources available for anyone who wants to take this further. Within this section we have provided links to some of the best resources we have come across, for example:

- General evaluation advice / strategies
- Evaluating activities in specific environments, for example mass media communications or Citizen Science
- Clarifying your intended outcomes
- Additional suggested tools and approaches, especially online & social media analytics and more detailed qualitative research techniques

2 – INTRODUCTION

This section provides a brief introduction to the European Evaluation Toolkit and its main purposes.

SECTION CONTENTS

2.1 — WHAT IS EVALUATION?

Evaluation is a systematic process designed to help you better understand - and improve! - your participants' experience of your outreach activities. It involves collecting information (for example feedback, observations or quiz results) and reflecting on what worked well, what could be improved, and what changed for the people involved as a result of their participation.

Based on feedback from members of the Europlanet community during the development of this toolkit, we have focused mainly on formative and summative evaluation techniques:

- Formative evaluation occurs whilst a project is still happening, particularly towards the start, and is designed to identify actions that can be taken to further improve the project before it ends.
- Summative evaluation can occur at any stage of a project, but is normally focused towards the end. It is designed to provide a summary of the achievements (results) of the project, based on the evaluation data collected.

The data collection tools we've included are well suited for summative evaluation (e.g. at the end of an activity or event), but can also be useful in formative evaluation (e.g. informing the development of future activities).

2.2 — WHY EVALUATE?

This is a question we're often asked! So, as part of the development of this toolkit we spoke to people in the planetary science community about the reasons that they evaluated their outreach activities. The main reasons identified fell within four key areas:

- **Improving** - Gathering feedback on participants' experiences to help identify what could be changed within the activity in order to improve it.
- **Providing evidence of knowledge gained** - Many outreach activities are specifically designed to im prove the audience's factual understanding of a specific topic (for example space exploration or exoplanets or astronomy more generally). As a result, there is a lot of interest in using evaluation techniques to find out exactly what participants learn from an activity.
- **Identifying other successes** - Some outreach providers are interested in other types of changes that occur for participants, for example whether they are inspired to follow a career in space science. We have provided links to a great framework, the Generic Learning Outcomes, to help people clarify what sorts of wider outcomes they are hoping to achieve. The final columns in the list of tools also explicitly indicate what sorts of tools best suit the main categories of outcomes.
- **Celebrating** - Providing an opportunity for both participants and outreach providers to creatively share the joy and inspiration, and for such excitement to be shared more widely (for example with funders and managers).

These identified purposes have provided the basis for the choice of tools showcased within the Europlanet Evaluation Toolkit.
We have also heard arguments that evaluations are pointless or 'only tell me what I already know'. In those cases it's often the funder that requires the evaluation evidence - so one important improvement could be to identify the most efficient / cost effective way to gather such data. Alternatively, you might be interested in a different approach that digs a little deeper and explores complementary aspects that you're not so sure about. For this reason we've tried to include a wide range of tools and techniques to evaluate outreach activities, to suit different needs.

2.3 — WHAT IS THE EUROPLANET EVALUATION TOOLKIT?

As part of the Europlanet 2020 Research Infrastructure (RI) project, funded by the European Commission under Horizon 2020, a dedicated outreach Evaluation Toolkit was commissioned in order to empower outreach providers and educators in measuring and appraising the impact of their activities. This toolkit is intended to provide advice and resources that can be simply and easily integrated into normal outreach and education activities.

2.4 — WHO IS THE TOOLKIT FOR?

This Toolkit is specifically designed to be used by outreach providers, educators and scientists who are interested in evaluating their outreach activities. Based on feedback during its development and piloting, this Toolkit does not assume any previous knowledge of evaluation. Instead, all the information and advice (including worked examples where possible!) is included to help get you started.

We know it might seem simpler just to provide a list of questions that you 'should' ask, or a set of downloadable forms that you can simply print and use (perhaps after translating them). However, because space science outreach covers a wide range of activities and topics, it is unlikely that any form or question we came up with would provide you with any information that is actually useful. Instead, we have tried to provide enough information, detail and examples to make it as easy as possible for you to conduct evaluation when you need to.

We are also not saying you have to evaluate *all* your outreach and education work! But if you do decide to do some evaluation - or are informed that you *must* evaluate something - we want to help ensure it goes as smoothly as possible. This toolkit is designed to help you with that process - to help you conduct evaluations that do not take too much time for either your audience or you, and that produce useful information.

2.5 — HOW WAS THE TOOLKIT DEVELOPED?

Professional outreach evaluators Karen Bultitude and Jennifer DeWitt (UCL, UK) were tasked with developing the toolkit, with strong support from Anita Heward, the Europlanet Outreach Coordinator. Initially, a series
of focus groups and scoping discussions were held with members of the planetary science community that are active in outreach in order to determine what they wanted from such a toolkit, and what sort of tools would be of most interest. A shortlist of tools was developed based on these discussions, with volunteers testing out the tool instructions once they were drafted. For some of the more in-depth tools the volunteers' experiences were developed into worked examples to help provide more concrete information about how to apply the tools in real-life situations. The Toolkit was also distributed to expert evaluators from the wider evaluation and public engagement communities for comment and feedback. Once the final revisions were incorporated into the Toolkit it was released publicly to the Europlanet community (and beyond!) for implementation.

2.6 — SPECIAL THANKS!

We would like extend a very warm thanks to all of the various wonderful people who have been so generous with their time and ideas during the development of the Europlanet Evaluation Toolkit. We also gratefully acknowledge the various stimulating collaborations we've had with many other different evaluation experts over the years, whose insights, rigour and expertise have creatively informed our professional evaluation practice, as well as the development of this toolkit.

In addition, we have highlighted within the specific tools where an individual or organisation has been closely associated with its development.

3 – STEPS TO CHOOSING THE RIGHT TOOLS

This crucial advice takes the form of a series of questions which will help you design your approach, and especially make your evaluation as efficient and effective as possible within limited time and resources.

SECTION CONTENTS

3.1 – WHAT DO I WANT TO FIND OUT?

Before starting any evaluation, it is important to have a clear idea about what it is you want to know, in as much detail as possible. This is similar to developing a research question. Defining what you're aiming to achieve from the evaluation at the start of the project really helps you focus your efforts, and identify what you do (and don't!) need to include. It also helps you avoid running into problems later on in the project. For example, a question like 'was this activity good?' is much more challenging to answer (what do you mean by 'good'?; for whom?) than 'Did most of the participants recognise the term 'black hole' by the end of the activity?'.

What you want to find out will determine the rest of the evaluation, especially the tools you use. For example, one very common evaluation question is whether an activity met its intended outcomes - click on the link for further information about how to do this effectively.

3.2 – HOW WILL I USE THAT INFORMATION?

It is often helpful to consider how you'll use the evaluation findings before you start. Do you need to convince stakeholders of the value of the activity? Feedback to funders? Make improvements in the activity? If you're primarily looking to make improvements, you might want to collect data that highlights which elements of an activity the participants found most engaging, or confusing. If you need to tell funders about who attended, you'll need to gather demographic information. If you need to convince stakeholders of value, consider having a conversation with them in advance about the kind of data they might find convincing.

Many academics often ask if it is possible to publish their evaluation findings. The answer is yes! However, to ensure the results are sufficiently rigorous for academic publication you need to approach the task with all the detail and research background that you would for a normal academic publication. For example, your methodology and approach would need to be appropriately informed by existing published work, and your audience recruitment and sampling at a level appropriate to the journal you intend to publish in. Deciding from the start that you hope to publish the findings therefore changes the level of depth and rigour that
you need to take in developing your evaluation approach.

3.3 – WHAT RESOURCE (TIME/BUDGET) DO I HAVE AVAILABLE?

It is important to be realistic when thinking about evaluation. Developing the evaluation tools, collecting data and analysis always take longer than you expect, especially when doing it for the first time. It is often better to set modest expectations (e.g. about how many people you might collect data from) and plan for high quality but on a small scale rather than collect data you don't have time to analyse or data that is so messy
it can't help you answer your questions.

Even if you can only collect limited data (which happens a lot in evaluation), you can still gain useful insights from it! You can also make the most of limited resources by partnering with others and by modifying or re-using tools others have developed or you have used previously.

At the other end of the scale, for larger projects you may wish to consider hiring an external evaluator to act as an independent 'critical friend' for the project. Participants are often more honest when talking to someone they perceive as separate to the activity, and external evaluators may spot things (both positive and negative) that would not normally be noticed by a member of the project team.

3.4 – HOW WILL I ANALYSE THE DATA THAT I COLLECT?

Again, it's worth being realistic here, and thinking carefully about what skills and resources you have access to. If you've never done statistical analysis or thematic coding before then it could really help to review how those processes work before you conduct your evaluation, to be sure that they will provide what you want. For example, if you're hoping to apply statistical analysis then that might affect the number of participants you need to collect data from (and how they are recruited). If you're planning to apply a 'light touch'

analysis such as word clouds then that might mean you don't need to collect demographic information. It might also suggest that you need to account for the time it will take to enter any handwritten data for computer analysis (see the "Resource" section above).

3.5 – SHOULD I CONSIDER WORKING IN PARTNERSHIP WITH OTHERS?

Working with others is a great way to make the most of limited resources. For example, you could ask a colleague to distribute questionnaires or conduct snapshot interviews at your event and then return the favour. If there are others conducting similar activities to yours, you could work together to develop common evaluation tools and even collaborate on the analysis.

You may also want to consider whether what you are doing could be part of a larger project. Are there teachers, researchers, students or others who might be interested in it as a case study activity? Are there funders or other organisations who might be interested in your approach and/or the evidence you gather, and willing to share the costs of data collection and analysis? Thinking creatively at an early stage might enable a much larger mutually beneficial evaluation to be achieved.

3.6 – ARE THERE ANY ETHICAL ISSUES I NEED TO CONSIDER?

Although you will have considered many ethical issues in planning your activities (e.g. around safety, causing emotional upset and so forth), there are additional things to think about when planning your evaluation.
In particular, collecting personally identifiable data (e.g. names, addresses, etc) can be very problematic in most countries for data protection reasons - so if you do not need to collect such data, then don't. Opt for an anonymous approach instead, perhaps using categories (e.g. postal codes) rather than uniquely identifiable specifics. If you do need to collect personal information, be sure to inform participants how that information will be used and stored, and get their explicit consent for such uses.

Demographic questions (e.g. gender, age, race/ethnicity, social class data) can also make people uncomfortable or even cause offense. So unless you need this information and have a clear idea of why and how you will use it, it is again best not to collect it. If you do need to collect it, word the questions sensitively (e.g. asking people to choose which age bracket they fall into, rather than provide an exact age), perhaps using questions from previous evaluations and/or asking for other people's opinions in reviewing questions that might be uncomfortable.

Working with children is also a common ethical consideration, as they can't legally give their own consent for participation. Sometimes it is possible for the teacher (or the school) to give permission for you to distribute a questionnaire or speak to certain individuals, but more commonly you will need to seek parent/carer permission for their child to be involved. This obviously takes time and advance planning (as it needs to be done in advance of the event), but many schools will incorporate it into the process when they seek permission for the children to attend the event, so if you speak to the school or parents in advance then you can work out how best to proceed.

It is probably obvious but it is also important to be aware that no-one can be forced to participate in evaluation! You need to give people the opportunity to opt-out, whether by not completing a questionnaire, politely declining a request for interview, or being excluded from any observations that might take place. Fortunately this is a rare occurrence (especially if you've made it clear in advance how the data will be used and why you're interested in collecting it), but do think about how you'll handle the process of gently seeking participants' permission to be involved, as well as what you'll do if they decline.
Finally, try to make your evaluations as short and enjoyable as possible. Not only will this increase the amount of data you can collect, but you don't want the evaluation to detract from engagement with your event or activity!

3.7 – WHAT METHODS WILL I USE?

Although this is often the first question that people jump to, we have good reason for including it last in our list here!! Until you have answered all the above questions it is easy to unintentionally select a data collection method that does not deliver what you want. For example, you might go to all the trouble of compiling and distributing a questionnaire, only to find that the results don't tell you very much, and it may

have been better to use an alternative such as snapshot interviews to get more specific, focused feedback. Worse still, your audience may have got bored or even frustrated by filling out "yet another" questionnaire, thereby detracting from their experience of an otherwise successful event.

Once you've answered the questions above, you're in a better position to decide what methods will be best for your activity or project. Do consider using multiple tools to gain different perspectives, or even as complementary stages in the process. For example, using 3 words at a pilot event with a smaller group might help you identify what questions to include in a pre-post quiz (perhaps using Mentimeter or similar) for later events with larger audience sizes.

3.8 — PATHWAYS TO SELECTING THE MOST APPROPRIATE TOOL(S)

To help you select which tool is most appropriate for your planned outreach activity we have prepared an overview table comparing their features. Based on feedback from members of the Europlanet community we have listed the tools in groups according to when they might best be used (during, beginning/end, or after an event).

Depending on your interests, you may find other pathways more helpful, for example, if you know your audience is 'Primary School' students then you can use that column in the overview table to help narrow down which tools are most appropriate. Other possible selection criteria that are included in the overview table (see the details worksheet) are:

- Suitable audience(s)
- Appropriate activity type(s) e.g. lecture presentation or drop-in workshop
- Response type e.g. online, written, verbal and whether it is short or long
- Time required to prepare the evaluation methods
- Time required to implement the evaluation methods
- Time required to analyse the data collected
- What sort of information is gained e.g. identifying what works and/or suggested improvements, evidencing changes that occured etc.
- Which of the Generic Learning Outcomes are most relevant to be investigated using each tool

The tools that contain case study examples from the Europlanet community are highlighted with **bold text** in the lists below and marked with an asterisk (*) and brief indication of the type of event the tool was used during in the lists below.

Tools best suited to use during an event

- Open palm on chest **(*public lecture)** - A simple, low-tech approach to gauge audience responses to multiple-choice questions during a lecture.
  — Similar technique: Traffic light
- Physical ranking scales - Participants are asked to physically stand along a line representing different levels of experience / attitude etc. (any ranking type question can be asked).
  — Similar technique: Yes, No, Maybe Runaround
- Graffiti wall **(*public lecture)** - Colour and artistic freedom combine to allow participants to respond to an event or activity in a highly creative way. [This could also be used at the end of an event.]
  — Similar techniques: Opinion postcards and Hand evaluation
- Mentimeter **(*teacher training session)** - A free online interactive presentation tool that allows presenters to quickly (and accurately) gather audience responses.
  — Similar techniques: Plickers, Poll Everywhere, and Twitter Clicker Alternative

Tools best suited to use at the beginning and/or end of an event or activity

- Snapshot interviews **(*school trip)** - Very brief, focused interviews, which are used in conjunction with an event to gather impressions quickly, like a photo of a moment in time.
- Pre/post quizzes **(*science café/public lecture)** - Brief surveys that are used before and after an event. They are ideal for helping you understand whether or not your audiences have learnt key aspects of the content you are trying to convey.
  — Similar techniques: Quirky questionnaires and Kahoots.
- Geographic location map **(*drop-in demonstrations)** - Where people come from can be really useful information, however isn't always easy to obtain. This simple and inexpensive technique encourag-

TOOLS OVERVIEW

	SUITABLE AUDIENCE(S)				APPROPRIATE ACTIVITY TYPE(S)						
TOOL NAME	PRIMARY SCHOOL	SECUNDARY SCHOOL	INTERESTED ADULT	GENERAL PUBLIC	DROP IN (festivals and demonstrations)	INTERACTIVE WORKSHOP	ONGOING SERIES (clubs, courses etc.)	LECTURE PRESENTATION	ONLINE	BRIEF DESCRIPTION	WORKED EXAMPLE INCLUDED?
Tools best suited to use during an event											
OPEN PALM ON CHEST	✓	✓	✓✓	✓✓				✓✓		A simple, low-tech approach to gauge audience responses to multiple-choice questions during a lecture	Image only
PHYSICAL RANKING SCALES	✓✓	✓✓	✓	✓		✓✓				Participants are asked to physically stand along a line representing different levels of experience / attitude etc. (any ranking type question can be asked).	
GRAFFITI WALL	✓✓	✓✓	✓	✓✓	✓✓	✓		✓		Colour and artistic freedom combine to allow participants to respond to an event in a highly creative way.	Image only
MENTIMETER	✓✓	✓✓	✓✓	✓✓		✓✓	✓✓	✓✓		A free online interactive presentation tool that allows presenters to quickly (& accurately) gather audience responses.	Yes
Tools best suited to use at the beggining and/or end of an event or activity											
SNAPSHOT INTERVIEWS	✓	✓	✓	✓✓	✓✓	✓		✓		Very brief, focused interviews, which are used in conjunction with an event to gather impressions quickly, like a photo of a moment in time.	Yes
PRE/ POST QUIZZES	✓✓	✓✓	✓✓	✓		✓✓	✓✓	✓	✓✓	Brief surveys that are used before & after an event. Ideal for helping you understand whether or not your audiences have learnt key aspects of the content you are trying to convey.	Yes
GEOGRAPHIC LOCATION MAP			✓✓	✓✓	✓✓			✓✓	✓	Where people come from can be really useful information, however isn't always easy to obtain. This simple and inexpensive technique encourages participants to provide their location data in a fun and visual way instead of a standard survey.	Yes
Tools best suited to use at the end of (or after) an event											
DROPPING, PEBBLES IN BOXES OR STICKERS ON CHARTS	✓✓	✓✓	✓	✓✓	✓✓	✓✓	✓			This process is quick, easy, and highly visual, and also provides an opportunity to participate for people who may find reading challenging.	
3 WORDS	✓	✓✓	✓✓	✓	✓	✓	✓	✓	✓	A quick, focused way to get a feel for participants' experiences by asking them to describe it in a few short words.	Yes
TARGET EVALUTION	✓	✓	✓	✓	✓✓	✓	✓	✓		A visual 'bullseye' approach to rating different elements of an event, or other outcomes of interest.	
POST-EVENT SURVEYS	✓✓	✓✓	✓✓	✓		✓	✓	✓	✓✓	Self-completion questionnaires that are used immediately after an event, workshop or programme.	Yes
PHOTOGRAPHY DIARY	✓	✓✓	✓	✓		✓	✓			Inviting individuals to select and discuss images that are especially meaningful to them, in their own words, thus providing insights that are otherwise very difficult to attain.	
PEER INTERVIEWS	✓	✓✓	✓			✓✓	✓✓	✓	✓	Peer interviews are a great way to encourage honest opinions, especially from teenagers, through involving participants in interviewing each other, and thereby hopefully allowing a more natural and honest conversation.	
TWEET SENTIMENT VISUALISATION		✓	✓	✓	✓	✓	✓	✓	✓	Automatic online analysis of tweets to help identify participants' reactions to a particular key word or phrase. Must be conducted within a week of the event (whilst the tweets are still "live" on Twitter).	

These characteristics are estimates based on extensive previous experience, however they may not hold true for evaluating every outreach activity.
See our dedicated advice to make sure the tool you choose is appropriate to your unique situation:
✓✓ = especially suitable to this category; ✓ = may be used

KEY TO SYMBOLS

This section summarises the symbols used within each of the tools to indicate where they are most appropriate.
Where a symbol is surrounded by a blue box within the tool text then that tool is especially suited to that particular category.

Suitable audiences (Who)

PRIMARY SCHOOL	
SECONDARY SCHOOL	
INTERESTED ADULT	
GENERAL PUBLIC	

Activity type (What)

DROP-IN (FESTIVALS AND DEMONSTRATIONS)	
INTERACTIVE WORKSHOP	
ONGOING SERIES (CLUBS, COURSES, ETC)	
LECTURE PRESENTATION	
ONLINE	

Data medium (Data)

ONLINE - LONG RESPONSES	
ONLINE - SHORT RESPONSES	
WRITTING - LONG RESPONSES	
WRITTING - SHORT RESPONSES	
MULTIPLE CHOICE RESPONSES	
VERBAL INPUT	
AUDIO / VIDEO	
DRAWING / VISUAL	

Time required (Time)

PREPARATION - SHORT (CAN DO IN HOUR BEFORE ACTIVITY)
PREPARATION - MEDIUM (DAY BEFORE)
PREPARATION - LONG (WEEK BEFORE)
RESPONDENT COMPLETION TIME (IMPLEMENTATION) - SHORT (<1 MIN)
RESPONDENT COMPLETION TIME (IMPLEMENTATION) -MEDIUM (1-5 MIN)
RESPONDENT COMPLETION TIME (IMPLEMENTATION) -LONG (>5 MIN)
ANALYSIS - SHORT (NONE NEEDED OR AUTOMATIC)
ANALYSIS - MEDIUM (HOUR/S)
ANALYSIS - LONG (DAY/S)

Information gained (Gain)

WHAT WORKS / IMPROVEMENTS
IMMEDIATE REACTIONS
EVIDENCE OF CHANGES THAT OCCURRED
MISCONCEPTIONS HELD

Generic Learning Outcome area(s)

KNOWLEDGE & UNDERSTANDING
ATTITUDES & VALUES
ENJOYMENT, INSPIRATION & CREATIVITY
SKILLS
BEHAVIOUR & PROGRESSION

We recognise that the above categories are not exhaustive, however from our initial scoping conversations with Europlanet members these were the major categories of interest for outreach efforts within Europlanet.

4 — TOOLS

— 3 Words

— Dropping pebbles in boxes

— Geographic location map

— Graffiti wall

— Mentimeter

— Open palm on chest

— Peer interviews

— Photograph Diary

— Physical ranking scales

— Post-event surveys

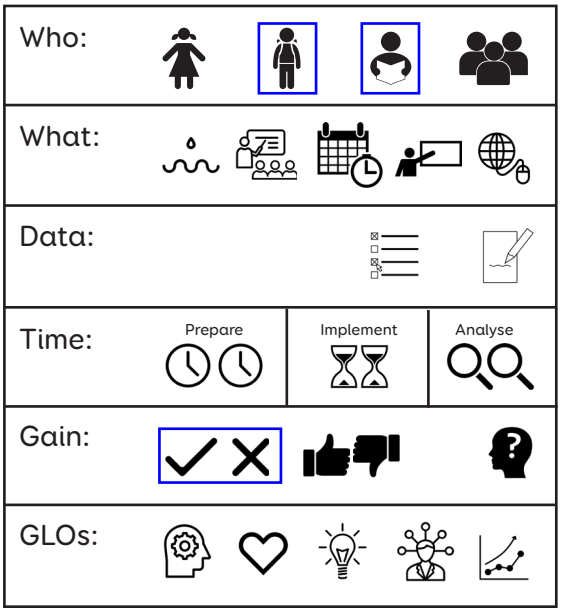
— Pre-post quizzes

— Snapshot Interviews

— Target evaluation

— Tweet Sentiment Visualisation

3 WORDS



A quick, focused way to get the feedback you need. Ask your participants to give you 3 words about your event (in priority order if you like). Choose your questions carefully and you'll be able to get not only a feel for their experience, but also focus on specific parts - what worked well, what could be improved, how it made them feel, and so on.

We've included instructions below, and there's also a case study example at the end to demonstrate how to successfully apply this technique.

WHAT DO I NEED?

- Post-it notes or scraps of paper - approximately 3 per audience member
- Pens/pencils (one per audience member)
- Box, wall display or other place for people to leave their responses once completed

Short written comments work really well, but this technique can also be done as part of a wider survey, or even a Twitter feedback stream (with an associated hashtag so you can track people's responses).

LET'S GET STARTED

1) Hand out the post-it notes and pens.
2) The simplest version is to simply ask participants to write down 3 words to describe the activity. But you can also be more specific, e.g.:

   What were the 3 best parts of today's event for you? (please label them 1,2,3 in priority order);

   What 3 things could we improve on for future activities?;

   What were the 3 most useful things we covered today?

You can also ask multiple questions - e.g. use different coloured post-it notes for responses to different questions, or simply get participants to label them (perhaps +/- symbols in a corner for what worked well / what could be improved).

3) Ask participants to place their responses on a wall display, in a box, or simply hand them in to a member of staff for collection once they are completed.

OK, WHAT DO I DO WITH MY DATA NOW?

There are lots of ways you can analyse these sorts of data. You can type the words into a list and create a word cloud to help visualise people's most common reactions (see our case study example below for a demonstration of how to do this with data collected using the 3 words technique). Or you can take a more systematic approach, grouping together different responses under common themes (this is covered (using the same dataset) in the thematic coding analysis outline).

GOT IT! HOW CAN I TAKE THIS FURTHER?

If you have time, and it's a manageable group size, get them to write each of their 3 responses on a different post-it note, then get the group to work together (e.g. using a large whiteboard or table) to go through all the responses to categorise them into common themes. This gives you great insights as to what participants naturally group together, any links they see between different elements, and a better understanding of their responses in the first place.

SIMILAR TECHNIQUES

Action plans or resolutions - Ask participants to each identify an action or change to their practice they plan to make as a result of the event or activity. Participants then take turns to verbally share their action plans or resolutions with the rest of the group.

Thought boxes - Everyone is asked to write down on a card their responses to a particular question of interest before the start of an activity and post it into a large decorated thought box (as part of an ice-breaker exercise). The facilitator takes out some cards and reads them aloud (anonymously) to start a discussion. The same question is asked at the very end of the activity, providing a before and after comparison of opinions. This technique can also be used to allow participants to ask questions they might be too shy to raise in person. It provides a great written record of their thoughts, and can incorporate drawing or other adaptations as appropriate.

Also see Graffiti Wall and other associated techniques.

3 WORDS EXAMPLE - WEBSITE FEEDBACK

THE ACTIVITY

Let's imagine we've developed a website, primarily for teachers to use, that collates useful information about relevant planetary science activities they can do with their students. For ease of discussion later we'll call this website Europlanet for Schools (EfS). The idea is that on EfS teachers can find exciting opportunities for in-class support, memorable days out, and engaging projects to motivate their students. Teachers can easily search for shows, workshops, debates, challenges, visiting speakers and more.

As part of the development of EfS we want to get teachers' insights into how it is all working, and where we need to make improvements to the site. The 3 words approach is especially useful in this case as it allows teachers to provide their reactions in their own words in a simple and effective manner - doing it online also means that it fits well with their likely involvement in the site itself.

Note that although this example is hypothetical it builds on existing data and good practice from elsewhere within the sector and is thus a realistic example of how to approach such an issue.

WHAT WE WOULD DO

A short pop-up online quiz asks teachers using the website to specify what 3 words they would use to describe EfS. The results are provided on the first worksheet (raw data) in the file qual coding. Using the approach described in word cloud analysis , we firstly create a simple visual map of the words provided by the 100+ teachers who we recruit to respond to the quiz (Figure 1). Within this picture, the larger the text the more common that word is within the dataset (word colours and placement are random).

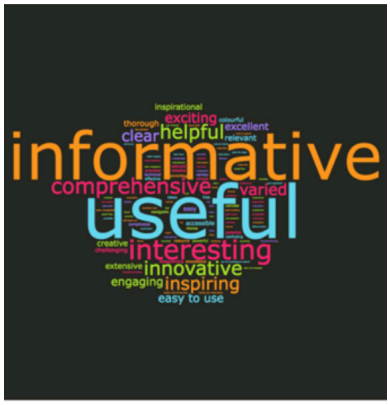


Figure 1 - Word cloud of hypothetical website feedback from 100+ teachers. See the word cloud tool for details regarding how it was produced.

The word cloud in Figure 1 is a quick and easy way to understand teachers' common reactions to the website. It is clear that the top two reactions from teachers are that it is 'informative' and 'useful', with many other positive words also being popular ('interesting', 'comprehensive', 'innovative', 'inspiring', 'helpful').

GOT IT! HOW CAN I TAKE THIS FURTHER?

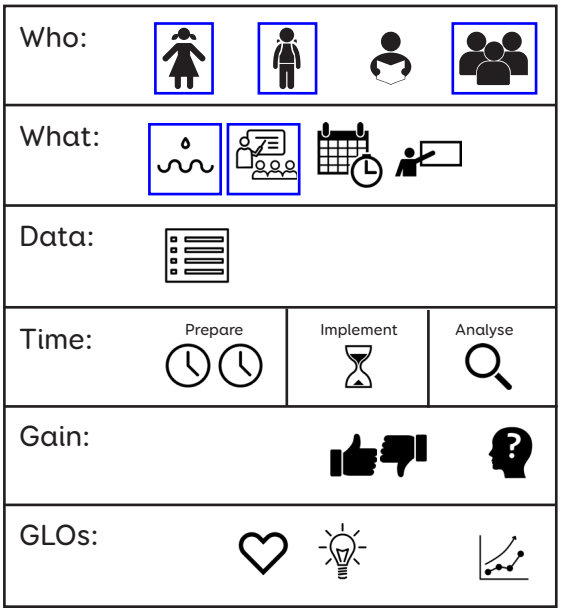
Further insights on people's reactions can be gained by grouping their individual responses into common categories. This process is the basis of all 'coding' of qualitative data - dividing it into common themes to help better understand the overall patterns in the responses. We've explored this further with the same dataset in our thematic coding tool.

WHAT WE COULD FIND OUT

Initially the image in Figure 1 can be used to help update the project team working on the Europlanet for Schools website, to quickly highlight teachers' main reactions. It could also be presented within an update report to the funder to help demonstrate the success of the current site, and justify further effort on it in the future.

Depending on the aims of the website we may want to tweak some of the content and then repeat this exercise in a few months' time to see if the responses match more closely to particular goals. For example, if one of the specific aims was to encourage teachers to use practical activities, then the current feedback in Figure 1 doesn't really include much mention of 'practical' elements. For greater emphasis, these could be highlighted more strongly within the website (perhaps with banners or specific news items), or linked to further from other popular resources, or more mention made of 'practical' components within existing content throughout many of the other pages. Then after a few months the same '3 words' exercise could be run to see whether these changes have had any influence on users' perspectives regarding whether 'practical' is a word they commonly associate with the site.

DROPPING PEBBLES IN BOXES



This process is quick, easy, and highly visual, and also provides an opportunity to participate for people who may find reading challenging.

WHAT DO I NEED?

- Pre-prepared container, for example a box, jar or chart.
- "Votes", for example pebbles, tokens or stickers.

Short written comments work really well, but this technique can also be done as part of a wider survey, or even a Twitter feedback stream (with an associated hashtag so you can track people's responses).

LET'S GET STARTED

Ask participants to provide a simple visual vote on a topic relating to the activity. Hand out your "votes" (usually one token or pebble per audience member) and ask them to place it in the box/jar or on the chart according to their opinion on the question asked. This usually works best if you have the question clearly displayed near the container, as well as mention it in person during the session so people are prepared.

OK, WHAT DO I DO WITH MY DATA NOW?

The beauty of this technique is that it's so simple to analyse. For many situations all you have to do is visually compare the levels of 'tokens' placed within the different boxes/jars. If you want to be a little more specific you can count the votes in each category, thereby allowing a more quantitative result.

This technique usually works best for questions related to attitudes or emotions, i.e. how people feel about or reacted to the event or its content. This broad view of the participant experiences can allow you to identify whether your activity is achieving what you intended, and photos of the final votes can be used for reporting purposes.

SIMILAR TECHNIQUES

There are many variations on this technique, for example:

- If you place two containers near the entrance/exit then separate votes can be taken: once as participants enter, and then again as they leave the event, to give you an indication as to whether any change has occurred. Be sure to indicate clearly which container is for use at the start and which at the end of their participation, and to keep the response options consistent so that you

can compare the before and after votes. Or you could use different coloured tokens in the same container(s) to track the votes before and after the event.

- Have multiple option slots in your container, so that participants vote from amongst a variety of possible options. For example, you could have a statement like:
  I want to follow a career in space science
  Then have options of Strongly agree, Agree, Neutral, Disagree, Strongly disagree for people to choose from, thus helping to judge participants' strength of feeling on the topic (rather than just a straight yes/no answer).
- Have multiple response options, such as different coloured stickers. For example, at the end of a series of sessions for an astronomy club a large sheet of paper could be set up and divided into the different sessions or areas of content covered. Participants could be asked to place different coloured stickers on the various sessions according to different knowledge or emotional reactions (Figure 1). Note that it is important to have a clear key, and to ensure each person only gets one of each colour sticker in order to avoid confusion! You can also invite people to write on the paper near their sticker to explain their vote, thereby providing slightly more in-depth information.

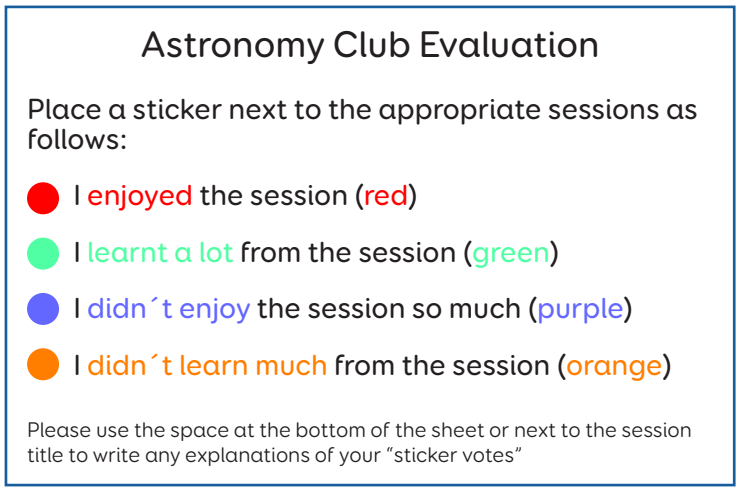
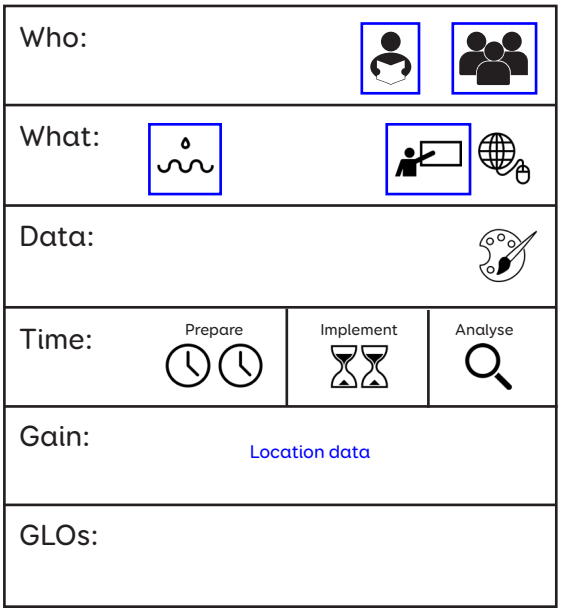


Figure 1 - Example of different sticky-dot voting options for an astronomy club evaluation

WARNING!! If using sticker-dot voting do be aware that there are some limitations - Dotmocracy has some great advice on how to overcome many of them.

- For younger children, use jars or plastic cups with options clearly written on them. Give each child a token (e.g. a piece of pasta or coloured card) to put in the jar that represents their preferred choice. It's usually best to cover the sides of the jars or plastic cups so that the children can't see others' responses. This means the children are more likely to think for themselves, instead of putting their token in the most popular option.
- For older audiences, use large see-through jars: many participants like seeing how their own response compares to others, and feeling "part of" the event in this way.

GEOGRAPHIC LOCATION MAP



Where people come from can be really useful information, however it isn't always easy to obtain. This simple and inexpensive technique encourages participants to provide their location data in a fun and visual way.

With thanks to Marianna Adams and the Isabella Stewart Gardner Museum for the original idea: https://artmuseumteaching.com/tag/measurement/

We've included instructions below, and there's also a case study example at the end which uses real data from the Europlanet community to demonstrate how to successfully apply this technique.

WHAT DO I NEED?

- Pre-prepared large-scale map
- Sticker dots or markers

LET'S GET STARTED

Create a large map of the area relevant to your participants - this might be a neighbourhood, city, country, region, or even the world, depending on what you are doing! The important thing is that the map needs to be sufficiently detailed so that it's easy for participants to identify 'their' part within the map. Place the map securely in an accessible area of your event - perhaps at the entrance/exit, or near where food and drink is served. Give each person a small coloured sticker dot and invite them to place it on the map where they come from (see the case study example below for images of this tool in action). It becomes a fun, social activity and, for some reason, people like to find themselves on a map, so it encourages participation from those who wouldn't normally provide such information.

You could also repeat this technique for other similar events and compare the maps - are people from similar backgrounds attending each of your events, or are there differences?

OK, WHAT DO I DO WITH MY DATA NOW?

You now have a visual picture of the geographic distribution of your audience. You can compare this to your intended audience and reflect its makeup: are you reaching the people you want to reach; are there any places which are over- or under-represented? A photograph of the distribution also provides a great visual display for any reporting associated with the event.

GEOGRAPHIC LOCATION MAP: EXAMPLE

With thanks to Anita Heward for contributing this example

THE EVENT

'Cornwall Sea to Stars' showcases the astronomy and geophysics intrinsic to and evident throughout the county of Cornwall, in the UK. The event described here involved a large trailer parked on a public quay over the Easter long weekend. Activities showcased included:

- Solar telescopes
- Binoculars trained on the Cathedral Spire
- Example telescopes
- A badge maker
- A 'Hunt the Planet' trail, with "planets" (laminated pictures of planets and information on the radius and comparison to Earth's radius) hidden around the quay (i.e. secured with plastic ties to benches, posts and other infrastructure).
- Handling samples of meteorites
- Videos
- Globes of planets and moons
- Spectroscopes and light-sources

WHAT WE DID

This event was the launch event for the 'Cornwall Sea to Stars' project, with a key aim to reach a diverse range of communities. It was therefore important to us to identify where people who visited the trailer came from. We mounted 3 maps on a large whiteboard, along with some coloured dots. The maps were each of a different level of detail: Cornwall only (with key cities marked to help those who were less familiar with map-reading); the UK and Ireland; and the world (Figure 1).

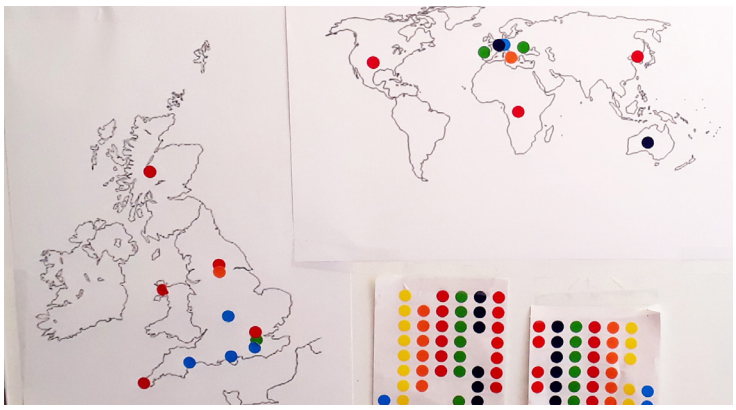


Figure 1 - Our three geographic maps at different levels of detail part-way through the event: Cornwall (top left), the UK and Ireland (bottom left) and the world (bottom right).

WHAT WE FOUND OUT

Of the (approximately) 500 people who visited the trailer over the two days, 49 people stuck a sticker on the map. This was lower than we had hoped, but likely due to the location of the board (see 'Reflections' below). Based on these data we could estimate a rough distribution of where our visitors came from (Figure 2).

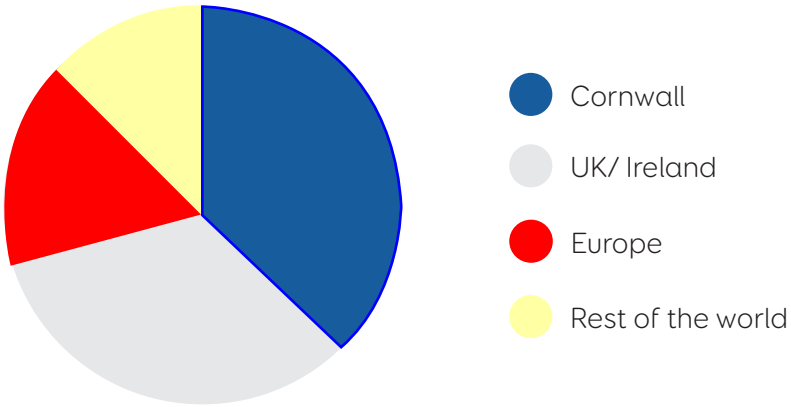


Figure 2 - Summary of where visitors came from

REFLECTIONS ON USING THE GEOGRAPHIC LOCATION MAP

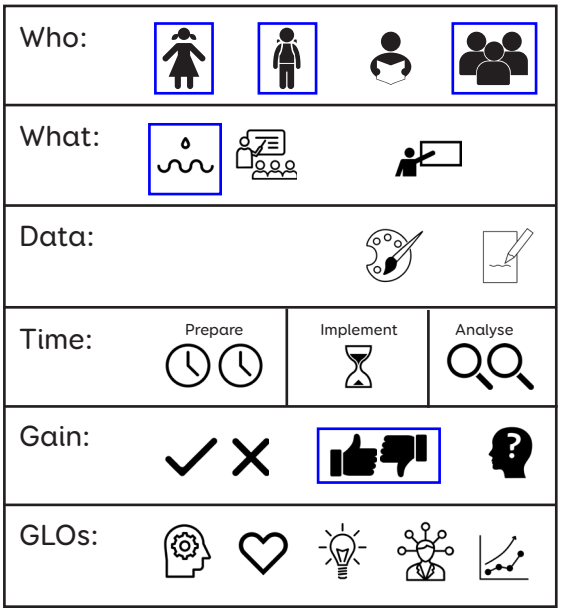
This was the most successful evaluation tool that we tested over the two days.

We had hoped to trial moving the whiteboard to a few different locations to optimise the positioning to get the most feedback from visitors. However, due to concerns about the strong winds that were present, the evaluation board was positioned in the shelter of one of the wing panels of the trailer for the majority of the two days.

It is probably unlikely that the distribution of stickers gives a true reflection of the breakdown of visitors from the local area vs further afield. The map in itself was something of a draw, with some people approaching the stand specifically to look at the map. People that had come from abroad appeared to be more willing/interested in sticking a dot on the map; we need to find ways to encourage more locals to place their dots - perhaps by running a tally competition between local towns / cities, or providing a more detailed map of specific areas of Cornwall?

Evaluation is not easy in a drop-in environment and there are physical limitations on where/how it can be conducted, especially in poor weather conditions, where writing/filling in questionnaires is impractical. Stickers seemed to work well and are relatively weather-resistant. The team will review strategy before the next event and adapt materials according to the conditions and audience anticipated.

GRAFFITI WALL



Colour and artistic freedom combine to allow participants to respond to an event or activity in a highly creative way. When done successfully, this approach can also create an attractive and eye-catching summary of participants' experiences that becomes an integral part of the event itself.

WHAT DO I NEED?

- Large wall space or board
- Large sheet of paper, post-it notes or other way of recording participant responses
- Prompt questions (optional)
- Coloured pens / markers

LET'S GET STARTED

Pin a large piece of paper onto the wall or board. Explain to participants (either verbally or in writing) that they are welcome to write or draw their comments, observations, feelings or messages (both to and from the group) to the graffiti wall at any time. Be as creative as you can - this will encourage your participants to do likewise: use different coloured post-it notes or pens, or even unusual surfaces linked into the theme of your event. You can allow participants to choose their own aspect to write about, or include some prompts for people to respond to. Some people find staring at a blank wall fairly daunting, so including a few questions or suggested topic areas can help to stimulate input from your audience (see Figure 1). If you think carefully in advance about these prompts they can also help structure the input received, thereby making it more useful for reporting purposes.

Figure 1 - Graffiti wall responses to an event run by Dr Yaël Nazé. Participants were asked to indicate what they liked (J'AI AIMÉ), what they didn't like (J'AI PAS AIMÉ) and any comments they had regarding the event (J'AI UNE REMARQUE). Note that at the time of the event these sheets were separated according to each of the questions asked, and of (roughly) equal size - their content has been combined here for easier visual display.

WARNING!! This technique does require a fair bit of proactivity in terms of encouraging your participants to get involved; it rarely works if you just leave the graffiti wall standing alone.

OK, WHAT DO I DO WITH MY DATA NOW?

This method creates a great visual representation of overall energy and engagement (so great for photos for final reports) however can be very challenging to analyse in depth! Many people find simply reading the comments helps them get a better feeling of participants' experiences, whilst others prefer to do a more structured analysis - see word clouds or thematic coding for details.

SIMILAR TECHNIQUES

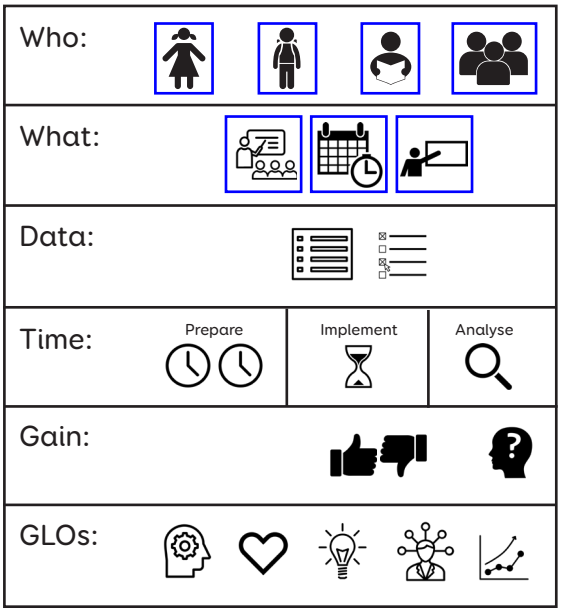
Opinion postcards - Postcards are provided with specific prompt questions for the respondents to complete. These may be specific (e.g. "What did you think of the event?") or more generalised / abstract (e.g. "I feel..."). Participants are encouraged to pin their responses onto a board or similar display for other people to see afterwards.

Hand evaluation - Give participants a piece of paper and access to pens. Ask participants to draw around their hand, and record the following on the fingers of their hand:

Thumb - something good, something they enjoyed
Index finger - something they would like to point out (could be good or bad)
Middle finger - something bad, something they did not enjoy
Ring finger - something they will treasure from the activity/event
Little finger - something little they want to add (could be good or bad)
Palm - A prediction for the future - What they are going to do next?

With thanks to the Woodcraft Folk for the 'Hand Evaluation' tool.

MENTIMETER



Mentimeter is a free, online, fun and interactive presentation tool that allows presenters to quickly (and accurately) gather audience responses. It can be used to collect a wide variety of data in an anonymous way. For example, presenters can use it to assess their audience's general level of knowledge in any specific topic using multiple choice quizzes or open questions, to allow audience members to ask open questions during a session, or to reflect on audience members' opinions about specific content covered, to name a few. Best of all, it allows the audience members to keep their responses confidential. It does require a reliable internet connection, and for each audience member to have access to a phone or computer to record each of their votes on the website.

https://www.mentimeter.com/

We've included instructions below, and there's also a case study example at the end which uses real data from the Europlanet community to demonstrate how to successfully apply this technique.

WHAT DO I NEED?

- A (free!) account on the Mentimeter website
- List of questions you want to ask - these are usually best as multiple choice options, though it is possible to ask short-response answers also
- Audience access to online devices (tablets, smart phones, computers etc.)
- Reliable internet access - for both the presenter and the audience to use
- (optional) A way to display the Mentimeter website to your audience - e.g. projection of your computer display

LET'S GET STARTED

There is plenty of support and information on the Mentimeter website, so we're only going to go through the basics here...
Step 1: Make sure the internet is working reliably on your computer.
Step 2: Sign into your Mentimeter account (free to register).
Step 3: Create your questions within the Mentimeter site - you can choose from many different question types including multiple choice or short-response, and you can also include images in your questions if you want.

Step 4: During your presentation, use the Mentimeter website to select the question you want to use. Once your audience has connected to the site using a special access key, they can provide their responses to your questions using their own phones/tablets/computers. The software will automatically count up the responses, providing you with instant feedback. You can also choose to display the results to your audience.

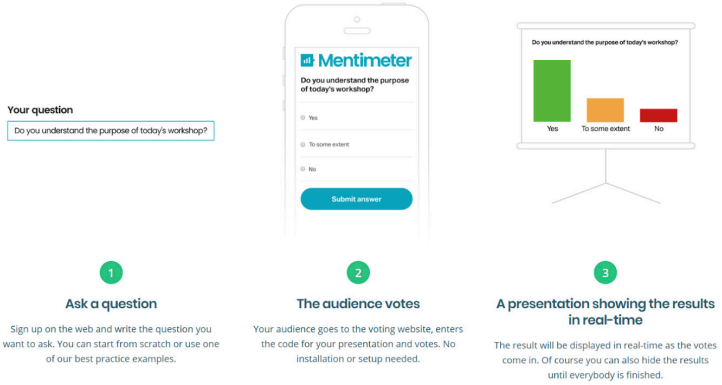


Image courtesy of https://www.mentimeter.com/

WARNING!! As with any computer-based technique, it's definitely worth checking the bandwidth and reliability of the internet connection at your event location well in advance. You should also confirm whether your participants will have access to smartphones, tablets or computers in order to complete the Mentimeter quizzes. For example, for smaller audiences can you take along a few spare devices in order to ensure everyone can participate (otherwise those who don't have a smartphone may feel left out)? We strongly recommend having a backup option ready!

OK, WHAT DO I DO WITH MY DATA NOW?

One of the great things about Mentimeter is that it automatically does all the analysis for you, so there's very little you need to do with it afterwards.

In the case of an ongoing series of activities (for example a weekly astronomy club), you could ask different questions each week, or even the same questions at the start and end of the series to see if individuals' responses have changed. If you ask each person to enter an 'identifier code' at the start (see the comparing responses section in pre-post quizzes for more detail on this), then within your Mentimeter account you can track their progress and see their results over time to each of the questions asked.

GOT IT! HOW CAN I TAKE THIS FURTHER?

Mentimeter is a great tool for helping ensure your content is relevant to the audience, at the right level, and that they understand important concepts before you move on. When used thoughtfully it can help you assess and react in real time to your audience's needs, making sure they get the most out of the session. For example, if the responses to a question you pose suggest that a lot of the audience don't fully understand something you just talked about, you could add in some further explanation and discussion of that content. If however that knowledge seems fairly well understood, then you can move on more quickly to new content.

SIMILAR ALTERNATIVES

Plickers - A free mobile app that allows presenters to run a multiple-choice vote and quickly (and accurately) count the audience responses using pre-printed voting cards, without relying on extensive technology.
Poll Everywhere - Live interactive audience participation during presentations
Twitter Clicker Alternative - Amalgamate responses in one place to a given question via social media using a dedicated hashtag, thereby giving your audience a free-response place to provide feedback or guess at a right answer.

With thanks to the University of Central Florida for the 'Twitter Clicker Alternative' tool.

MENTIMETER: EXAMPLE

With thanks to Priscila Doran for contributing this example, with input from Joana Latas and Rosa Doran

WHO WE ARE

NUCLIO is a non-profit organization and an NGO for development composed of professional scientists, researchers and teachers that focus on science outreach and teacher training in the framework of science education, and education development and innovation. NUCLIO is an official training centre recognized by the Portuguese Ministry of Education and is the coordinator of the Galileo Teacher Training Program, a training program that was born in 2009 and has already reached 50 000 teachers in over 100 nations worldwide.

OUR EVENT

One of the main goals of NUCLIO is to provide teachers with the tools to reach their students in a deep and meaningful way, promoting engagement and the need to learn. In the fast digitally developing 21st century, we consider that integrating digital technologies in everyday teaching is of tremendous importance. As such, instead of just presenting these tools to teachers, we focus on using them in the course of our training sessions in order to show their potential and help teachers to be comfortable in using them. In this example we focus on one of these tools: Mentimeter.
During a teacher training session within the framework of the project PLATON (http://platon-project.eu) the trainer used Mentimeter in the form of a pre-test / post-test regarding the teachers' perceptions of different components of Inquiry-Based Learning in their daily practice.

WHAT WE DID

When preparing the session, the trainer accessed https://www.mentimeter.com to write the questions to be presented to the teachers. At the beginning of the training session, the trainer asked the teachers to use their mobile phones to connect to https://www.menti.com and access the session questions using a specific code generated by Mentimeter (Figure 1).

Figure 1 - Participating teachers using their mobile phones to access the pre-session questions

Once each teacher entered the appropriate code the questions appeared one by one on their screens along with the answer options. Mentimeter provides a wide variety of answer options, including open response, scale, multiple choice, etc. In this training, each teacher was asked to use a scale to evaluate from 1 to 6 the extent to which extent they already used each of the components of Inquiry in their daily teaching practice (Figure 2). The answers were presented immediately in the trainer's screen, in an anonymous way.

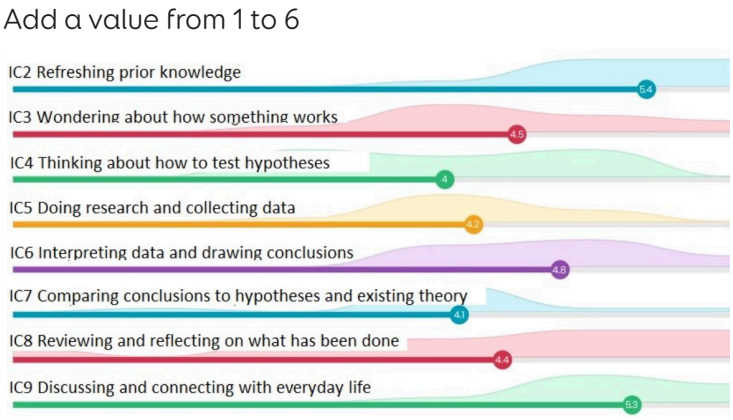


Figure 2 - Participants' responses to the pre-session quiz about teachers' use of Inquiry in their daily teaching practice

After this pre-session quiz the trainer introduced each Inquiry component in depth to the teachers. By the end of the session, the same process was repeated using Mentimeter, and teachers again evaluated their use of each of the components, now understanding their meaning more in depth (post-test) (Figure 3).

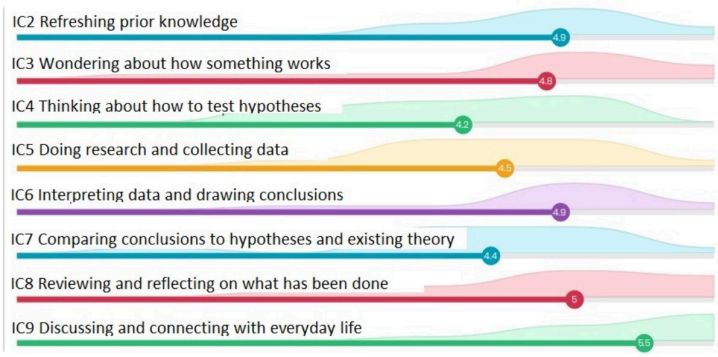


Figure 3 - Participants' responses to the post-session quiz about teachers' use of Inquiry in their daily teaching practice

Finally, at the end of the session, teachers were asked to access the platform one last time and describe the training session in 1 word (Figure 4). Here again, Mentimeter provided a variety of ways to present these words: list, grid, one by one, word cloud, etc. In this training the trainer chose to present a word cloud where the most common words appear in the centre in a bigger size than the others (see the word cloud analysis technique for more information on this technique generally, though it is calculated automatically within the Mentimeter software).

Figure 4 - Word clouds of 4 training sessions in the framework of PLATON, composed of teachers' choices of 1 word to describe the session they attended. Language: Portuguese

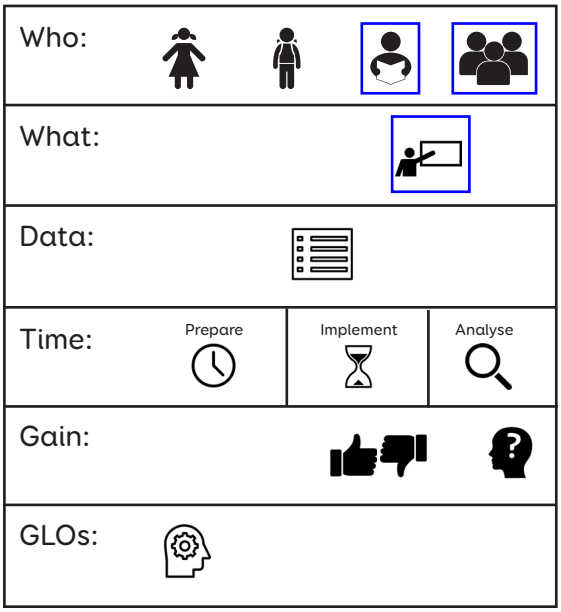
WHAT WE FOUND OUT

In this session, the use of Mentimeter presented valuable results (Figures 2 and 3) as it showed the trainer that most teachers felt they already used Inquiry practices in their daily teaching, and that some components were more widely used than others. This was an important starting point in identifying which

components of Inquiry the trainer should focus on in more depth than the others for that particular training session.
In a wider context, compiling the results from multiple sessions will provide valuable information of how teachers interact with each Inquiry component in their practice, allowing for the adaptation of the training program as a whole. Also, comparison between figures 2 and 3 demonstrates that the answers change from the pre-test to the post-test. This allowed the trainer to understand that Inquiry was not a clear methodology to some teachers and that dividing it in different components and training teachers deeply in each of them was a very valuable way of promoting teacher's knowledge, confidence and motivation in implementing Inquiry as a whole in their practice.

The word cloud results allowed the trainers to understand that the participating teachers saw the PLATON training session as mainly interdisciplinary, innovative and interesting, but also motivating and challenging.

OPEN PALM ON CHEST



With thanks to Dr Yaël Nazé for suggesting this technique

A simple, low-tech approach to gauge audience responses to multiple-choice questions during a lecture. It also has the added advantage of keeping respondents' answers relatively private (compared to more standard "hands up" techniques), and being applicable to audiences of all ages.

WHAT DO I NEED?

Nothing but your audience and a clear view of them - presenting at the front of a lecture theatre is perfect.

LET'S GET STARTED

Ask your audience members to each hold one hand close to their chest in response to a question you pose. An open hand means "yes" and a closed hand means "no" (or true/false if you prefer).

[Alternatively, you could just ask people who agree with a statement to hold one hand open on their chest, as in the picture here].

For example, you could ask:
Does the Sun orbit the Earth?

Looking around the room you can quickly see (roughly) what proportion of people have open palms compared to closed palms, and therefore better understand your audience members' knowledge. Holding their hands close to their chests also "feels less like school" and avoids the problem of influencing neighbours' responses.

OK, WHAT DO I DO WITH MY DATA NOW?

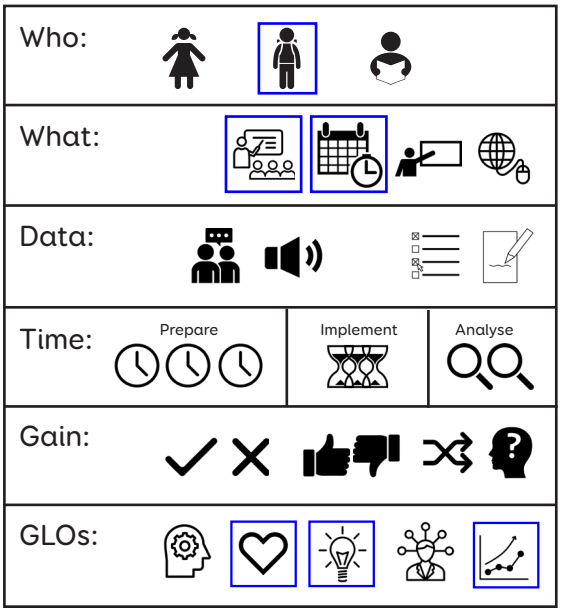
Asking simple and relevant knowledge questions within your presentations will allow you to better gauge the understanding and backgrounds of your audience, and therefore adjust your content appropriately. For example, if there were a lot of "no" responses to the above question then you might need to start with some basic facts about the solar system, however if the majority of people say "yes" then you can move on to other content more quickly.

SIMILAR TECHNIQUES

Traffic light - Participants display coloured cards according to their responses / confidence / agreement relating to a posed question.

With thanks to Rochester Community Schools for this technique.

PEER INTERVIEWS



Getting honest opinions from participants can be challenging, especially with teenagers or other groups where the evaluator is seen as an authority figure. This can lead to participants adjusting their responses according to what they think the interviewer does (or sometimes does not!) want to hear. Peer interviews are a great way to overcome such problems, involving participants interviewing each other, and thereby hopefully allowing a more natural and honest conversation.

WHAT DO I NEED?

- Pre-prepared prompt questions or topic guide
- Mechanism to collect responses (audio recorder(s), paper, follow-up discussion or similar)
- Time and space to hold the interviews

LET'S GET STARTED

Effective briefing is essential for this technique. It may be helpful to create written instructions for participants to follow, or if time allows, have a short practice attempt and debrief prior to the "real" peer interviews. The idea is that the participants take turns in the roles of "interviewer" and "interviewee", working with other participants to reflect on the event or activity.

Start by agreeing on the questions to be asked. You may wish to prepare this yourself in advance (and distribute a fixed list of prompt questions to your interviewers), or alternatively you can identify broad categories of interest, and ask the interviewers to come up with their own questions. The latter approach is more time consuming, but does have the advantage of authenticity, and having the questions phrased in their own natural language.

It's also important to think in advance about how the data will be collected and subsequently analysed. For example, you could use audio recorders (modern smartphones do the job very well), though this involves challenges in transferring the data to a central location, and the time required to listen to all the interviews in detail. Alternatively, you can ask your interviewers to write a short summary of the responses to each of the questions they ask, or simply hold a general discussion with all of the interviewers afterwards, asking them to reflect as a group on the responses to each area of interest.

Once you know how you will collect and analyse your data, brief your interviewers, and then provide them with space and time to go through those questions with other participants. Depending on the group involved this works well as either a paired or small group exercise, i.e. one interviewer speaking to another peer, or to a small group of their peers to gather their impressions.

Don't be afraid of involving your participants (especially teenagers) in many of the above decisions. Being part of the process, and having an input into how it is run, can be very empowering, and often means that they take it more seriously.

OK, WHAT DO I DO WITH MY DATA NOW?

Peer interview data is similar to any other form of qualitative data, and can be analysed in either a light-touch or an in-depth manner depending on what you want to achieve, and what you intend to do with the information collected. For example, a group discussion after the peer interviews are conducted might allow you to identify particular strengths and weaknesses of the event that you can then incorporate into future planning. Alternatively, if you ask your interviewers to either audio record or write their responses, then you can analyse the written text (including transcripts of audio data) using the thematic coding approach, and use direct quotes to highlight key themes within your reports. This more in-depth analysis can then be included in reports or even future funding applications to reflect key outcomes of your programme, and convince others to support it further in future. The authentic "participant voice" that results from peer interview techniques can be very powerful in convincing wider stakeholders of the importance and/or effectiveness of a programme.

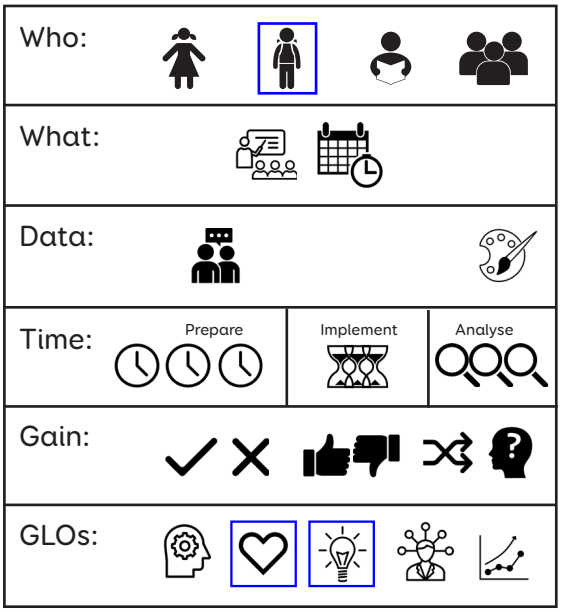
SIMILAR TECHNIQUES

Vox Pop - Create a set of evaluation questions for the young people to ask each other. Using simple video equipment such as a mobile phone, ask the young people to record the interviews. If the young people are uncomfortable with being filmed then you can film away from their face and film hands or feet while they're talking instead. It's very important to make sure the audio is recorded well. Also highlight that the film will only be used for evaluation and no other purpose.
Psychoanalysis - Participants get into pairs and interview one another about the event or activity. Importantly, the focus is on analysis of their experience and attitude towards it rather than factual recall of its content. For example, typical prompt questions could include:

Can you describe to me the event/activity that you experienced today?
What were your attitudes/beliefs before this event?
How did your attitudes/beliefs change after participating in this event?
How will/have your actions/decisions altered based on your participation in this event?
How have your perceptions of others/events changed?

With thanks to the Woodcraft Folk for the 'Vox Pop' tool and University of Central Florida for the 'Psychoanalysis' tool.

PHOTOGRAPH DIARY



Images (especially photos) can help capture the richness and depth of your participants' experiences. In particular, inviting individuals to select and discuss images that are especially meaningful to them, in their own words, provides great insights that are otherwise very difficult to attain. It is particularly helpful during the implementation of a new programme of activity, to check how things are going and provide formative feedback opportunities.

There are quite a few different ways you can implement this technique, but here we talk you through one of the simplest / easiest to deliver.

WHAT DO I NEED?

Ability to capture, collate and view photos from your event. Depending on the resources available (and security situation!) you could for example:

- Invite participants to take their own pictures and upload them to a Facebook, Twitter or Instagram feed using a designated hashtag.
- Distribute digital cameras or tablets to groups (up to 5 or 6 participants per group), with instructions for the group to take photos of anything particularly special or meaningful from their experience of the event. Then collect the cameras/tablets at the end and download their photos to a shared folder.
- Take your own photos regularly throughout the event, trying to capture images of both the event itself as well as the participant 'experiences' of it (audience reactions & viewpoints).

People often need encouragement to actively take and/or publicly share their photos, so if you go for one of those approaches then make sure you include reminders (and possibly even set times during the event) to allow for this.

It can be really helpful to be able to handle the actual photos during your discussions, so if you have the facilities then try to print them out (the quality doesn't have to be perfect). If not then you'll need a computer with a reasonably large screen so you can view them during your discussions with the participants afterwards.

You'll also need a quietish place to be able to hold individual or small-group discussions with some of your participants after the event, where they can see the photos / computer screen. Allow at least 20-minutes for each discussion, more if you plan to go into further detail. You'll also need a way to invite people to participate in the photograph diary discussions - for example, as part of the event itself (e.g. if they are already working in small groups it could be their final activity), or to volunteer at the end of the event, or pre-arrange a small group to stay behind afterwards.

Finally, think in advance about how you plan to record your discussions - it can be helpful to have a separate person available to write down the details while you focus on interacting with the participants (or vice versa).

LET'S GET STARTED

If your photos are in hardcopy then you can probably handle 6-8 people per discussion. Depending on the level of depth you want to go into you may be able to run multiple separate discussion groups. If you are taking the computer approach then the group size is likely to need to be a little smaller (2-3 people maximum). To start, briefly welcome your participants and thank them for their time and input, and explain how you're going to use the information you collect today. For example, a typical approach could be:

Thank you so much for staying behind to speak to me now. Your input will really help me understand your experience of the event, so that we can improve our work in the future. Before we start I just want to emphasise that there are no right or wrong answers, so please just give me your honest opinion to any of the questions I ask. My colleague <<Name>> is here to help capture your responses so that I can concentrate on our discussions, but nothing you say will be attributed to you individually in any reporting. If anything is unclear at any stage please do feel free to ask questions.

Once your participants are comfortable and happy to proceed, display the photos randomly on the table / computer screen and ask participants to individually select one that is particularly important to them. You may wish to use specific prompt questions here relating to your intended outcomes for the event, for example:

- *Which image best represents your experience of the event?*
- *Which image best represents the most memorable part of the event for you?*
- *Which image is best linked to your favourite part of the event?*
- *Which image is best linked to something you feel could be improved about the event?*
- *Which image is best linked to a point during the event where you learned something new?*
- *Which image is best linked to a time where you had an emotive reaction during the event? For example, you felt inspired, surprised, angry, frustrated, or enjoyed yourself?*
- *Which image best represents the part of the event that you are most likely to tell someone else about?*

Note that you would select 2-3 questions maximum (maybe more for smaller, computer-based discussions), or the process is likely to start getting boring for everyone involved! Normally it's better to ask each question in turn (to reduce confusion), but if you're pushed for time you could combine them, e.g. ask participants to select say 3 separate images in one go, representing e.g. 1) their favourite part, 2) a point where they learned something, and 3) something they felt could be improved.

Emphasise to your participants that they should go with their initial instincts regarding which image(s) they select (so not to over-think their choice), and also that they should then try to think briefly about the reasons that prompted that choice. Encourage your participants to physically pick up their chosen image if possible, then go around each person in the group in turn, inviting them to share which image they have selected, and why. You may need to prompt those discussions to delve a little further e.g. "What was it about that demonstration that you really liked?", or "Would you mind describing what it was that made you feel < <selected emotion> > at that point?" When recording the details of these discussions, make sure you capture which image related to each explanation (and if not obvious, which particular part of the event the participant felt it related to).

Once everyone has had a chance to go through their image(s), try to summarise any common themes or contrasting experiences, and invite wider discussion (e.g. responses to other people's points). It's important to let the participants tell the story in their own words, and to come up with their own links - this will give you the best insights into any misconceptions or potential outcomes that your event has achieved.

(See our question types section for further information on designing prompt questions; we'd advise using open questions where possible within these discussions).

OK, WHAT DO I DO WITH MY DATA NOW?

Once you've finished your discussions then read through the notes and try to identify any common patterns, and what they mean for your programme. Are there particular parts of the event that regularly prompted audience reactions (whether positive or negative)? Do the participant reactions match what you hoped to achieve? Did any of the reactions surprise you? What can you change in response to the reactions recorded?

For most purposes, a reflective, thoughtful review of the data is likely to be sufficient, without further detailed analysis. However, if you wish you could also apply the thematic coding approach to the qualitative data you collect regarding participants' reactions. You could also extend this from being text-based to a visual analysis regarding the content and/or style of images selected in response to each question.

More broadly, think about how you might be able to report the reactions in a useful way. For example:

- Report to a funder: "In post-event small-group discussions with participants, <<activity>> was most commonly described as the point where they learned something new, thereby justifying our focus of resources on that element."
- Internal team meeting: "Here are some images chosen by participants as points that they felt we could improve. <<Briefly summarise participants' reactions to each image>>. Any thoughts on what we can do to overcome these issues?"
- Project publicity: "Interested in <<topic>>? Want to know more about <<subject>>? Come and be part of our next event, <<Name>>. Previous audiences have particularly loved our <<activity>>, and have reported <<insert statement about learning benefits / evidence of other outcomes>>."

If the participants have taken any particularly useful images that they've shared during this process you could also ask permission to use them, for example in reports, or on the programme website.

GOT IT! HOW CAN I TAKE THIS FURTHER?

The technique described above assumes that you use images of the event itself, however it can also be successfully applied using more generic images. This is particularly useful for conducting front-end or formative evaluation, to inform further planning / revision of the event or activity itself. Prepare a series of say 20 postcards / images representing a range of people / situations / backgrounds etc. The actual content is less important than ensuring there is good contrast between all the different images - for example, some with people (of different genders, ages, backgrounds in different locations and environments, and with different facial expressions), as well as images of animals, landscapes, buildings, constructions, weather... whatever you can source! The point here is not so much to ensure that the images are linked to your event, but more that they provide enough variation to stimulate different reactions and emotions within your participants. For example, whilst an image of a closed door might not seem relevant to an event on astronomy, a participant might choose it and say "I felt like the presenter was not interested in my questions about astronomy and was shutting me out". Your participants will draw their own links between the images and the emotions they experienced, but the important thing is to provide a wide variety of images to encourage a full exploration of participants' perceptions.

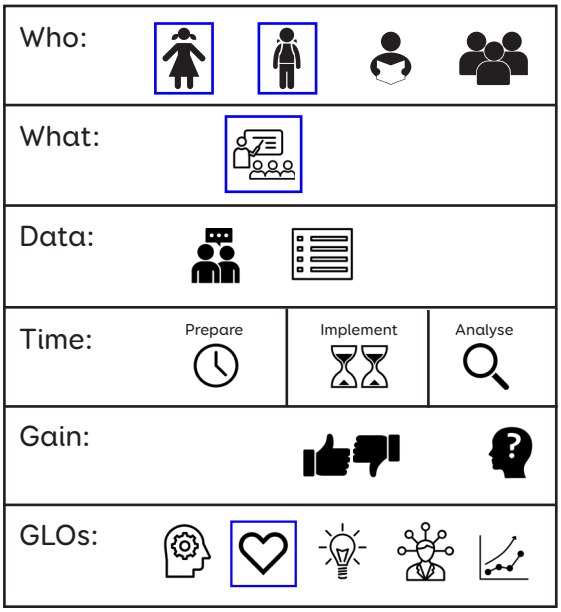
You can then run the discussion in the same way as described above, with the advantage that the cards can be pre-prepared (and possibly even numbered for easy identification during reporting). You do however lose the possibility of triggering reactions to specific parts of the event, as well as seeing what aspects participants themselves choose to take photos of.

SIMILAR ALTERNATIVES

Scrapbook Selection - Put students in groups and give each group a big pile of printed photos (best if laminated - maybe different shapes/sizes?) Ask them to choose one as a group that epitomizes their reaction/definition of the topic being discussed, and explain why.
Digital storytelling - A computer-based approach where participants put together a narrative based on their chosen images. For example there are existing lists of digital storytelling tools suitable within educational environments.
Individual photo interviews - You can also run the photo diary technique with individuals instead of groups.

With thanks to University of Central Florida for the 'Scrapbook Selection' tool and Learning Space Toolkit for the 'Individual Photo Interviews' technique.

PHYSICAL RANKING SCALES



Participants are asked to physically stand along a line representing different levels of experience / attitude etc. (any ranking type question can be asked). This gets everyone moving and raises the energy in the room, and also allows for interaction and discussion as the participants try to place themselves in relation to others around them. It is a very energetic approach, though it can be difficult to quantify.

WHAT DO I NEED?

Your participants and space in the room for them to move about!

LET'S GET STARTED

Explain to your participants what is going to happen before they start moving, otherwise things can get very chaotic very quickly! Indicate an imaginary scale within the room (between two walls is usually best, but anywhere with a reasonably long space is good). Ask your participants to physically arrange themselves along that line according to the question asked. For groups that are new to each other, but will be working together extensively, it's often useful to begin with a simple question that everyone can answer easily, for example "How far have you travelled to get here?", with a scale running from "closest" to "furthest away". You can then move to more complicated / in-depth questions, for example, the line might go from "*Expert*" to "*Beginner*", with the question "*How confident are you in programming skills?*". Or from "*Strongly agree*" to "*Strongly disagree*", with the question "*Do you believe there is life on other planets?*"

Encourage your participants to speak to each other in order to better judge where they should be on the line - it's great as an icebreaker at the start of the session to encourage people to get to know each other a little better, as well as providing information to you.

Once they are arranged in a rough line (you may need to hustle them to finish deciding where to stand) you can ask certain individuals (at different points along the line) to briefly share why they placed themselves where they did. Repeat this with a few other brief relevant questions - these could be topic related or more general. Depending how much discussion you include, around three questions is usually enough.

OK, WHAT DO I DO WITH MY DATA NOW?

The combination of the ranking scales with the brief discussions helps you understand participants' opinions and experiences, and allows you to respond to any misconceptions or comments that people might have. It

also helps participants understand who else is in the room, what sort of background/skills they have, and gives you a chance to emphasise that diversity is welcome, and that everyone has something to contribute.

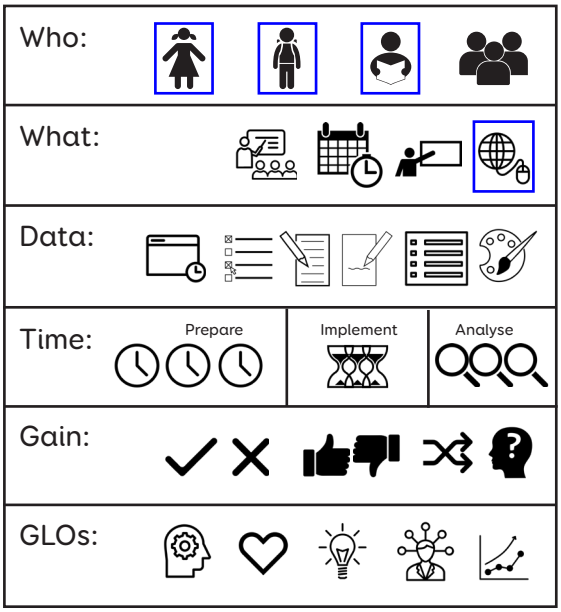
SIMILAR TECHNIQUES

You can repeat the exercise at the end of the session to help identify any shifts in attitudes or confidence (and even take photos of the group before and after for easier comparison if they give permission).

Yes, no, maybe runaround - This is an active way of collecting children's opinions. You will need three large pieces of paper with a smiley face on one, a sad face on the second and an unsure face on the third. Position the pieces of paper around the room and shout out your question. Ask children to run/hop/skip/walk/tip toe/jump etc to their preferred answer. You will need someone to record how many children went to yes, no, maybe. Add in some fun questions too!

With thanks to Sheffield City Council for the 'Yes, no, maybe runaround' technique.

POST-EVENT SURVEYS



Post-event surveys are self-completion questionnaires that are used immediately after an event, workshop or programme. They help you understand the experience of your audience - what they thought of the experience, what they feel they have gained from it, what they think could be improved. Their advantage is that they can be used to gather data relatively easily from large numbers of individuals, but they are not well-suited to gaining insight into the how's and why's, nor to gathering robust evidence of change (e.g. of demonstrable increased understanding). They may, with caution, used to ask participants about their own perceptions of what may have changed for them (e.g. 'What did you learn today?').

We've included instructions below, and there's also a case study example at the end which uses real data from the Europlanet community to demonstrate how to successfully apply this technique.

WHAT DO I NEED?

Surveys can be delivered on paper or online, if you have devices available. If you decide to use paper versions, you'll need to prepare plenty of copies in advance and have pencils or pens available at your event, as well as flat surfaces to write on. It's fine if clipboards aren't available as long as there are tables or other surfaces around.

When planning the event, make sure to have time set aside at the end when people can fill in the survey. In addition, while surveys can be straightforward to administer, response rate can be a problem. If at all possible, it is a good idea either to facilitate the surveys directly (ask the questions and fill in participants' responses) or to have a few people (colleagues if possible) hand out the surveys and be available to clarify questions and collect completed surveys.

LET'S GET STARTED

Prior to the event, you need to create the survey. See our Steps to choosing the right tools advice to help you do this effectively. For example, your survey items (questions) should relate directly to your aims for the activity you are evaluating and should, ideally, provide you with information you might find useful for improving your practice or feeding back to a funder. Because it is important to keep surveys as short as possible, it is best to leave off questions related to, for instance, background knowledge of the topic.

To start, consider what you want to find out from the survey. For instance, you might want to cover participants' overall experience of an activity (e.g. whether they enjoyed it, found it interesting or dull). You may also want to find out about their experience of specific elements of a broader activity (e.g. whether particular elements were clear or confusing, whether they felt they could contribute to a discussion) or whether they felt they learned about a particular topic or felt more confident in their understanding. You might also want to know about their motivations for participating (i.e. why they had decided to participate in a particular event) or whether their expectations for the event were met.

We've outlined the main types of questions you might want to consider using below; as you'll see in our case study example, most surveys combine multiple different question types in order to capture a broad range of information.

Question types

There are many different question types that can be used, but some of the most common types are (see the pre-post quiz section for some other question types):

Yes/No questions:

Have you been to the space expo before? Yes/No

Rating scales:

On a scale of 1 to 5, where 1 is not interested at all and 5 is very interested, how interested are you in going to another talk about supernovas?

1 2 3 4 5

How clear or confusing were various parts of today's talk?

	Very confusing	*Confusing*	*Neither*	*Clear*	*Very Clear*
Introduction					
Description Of Astronaut Selection Process					
Demonstration Of Weightlessness					

How much do you agree or disagree with the following statements?

	Strongly Disagree	*Disagree*	*Neither Agree nor disagree*	*Agree*	*Strongly Agree*
Today´s speaker was an expert					
I did not have enough opportunity to contribute to the discussion.					
I would like to find out more about this topic.					
I enjoyed this show.					
This event has increased my interest in this topic.					
After going to this event, I feel more confident to discuss space travel with others.					
My expectations for this event were met.					

Please note - the above statements are broad examples. It is best to tailor your questions to your specific event or elements of an activity, to make it as clear as possible what you're asking. In addition, it is very important that rating scale response options are balanced - that there are as many negative response options as positive (you'll see that the choices in the examples above are balanced). In contrast, the options 'really strongly agree, strongly agree, agree, neither, disagree' are skewed in a positive direction.

Open-ended questions:
These questions give your respondents a bit more flexibility in how they answer. However it is also important to keep them as short as possible to ensure they are clear. For instance:

What would be one thing you would change about the show?
What would be one thing you keep the same?

When creating open-ended questions, it is important to keep them as clear and concise as possible, or people will not respond. In addition, people often skip them anyway, so if you decide to include them, they should not be critical to what you are trying to find out.

Demographic questions:
You may also want to include demographic questions such as:

What is your gender?
MALE FEMALE OTHER PREFER NOT TO SAY

Which age bracket do you fall into?
18-24 25-39 40-54 55+

These kinds of questions will give you some insight into who attended the event and their experiences. However, because not everyone will respond to the questionnaire, they may not provide an accurate reflection of participants. In addition, these questions can be perceived as intrusive, so unless you need to ask them, you may want to leave them off. If you do decide to include them, it is advisable to keep them to a minimum.

Types of questions to avoid

Writing good questions – ones that are clear and concise and that provide data that will be useful – is not easy. Here are some example questions that might seem fine but are actually problematic:

Loaded or leading questions (where it is obvious what kind of answer you want):

Was this the most exciting show you've ever been to?
Yes/No

Are the public well informed about scientific developments or are scientists deliberately keeping them in the dark?

Planetary scientists often use most of their research funding to throw all night parties where they take lots of illegal drugs. Should the government increase research funding for planetary science?

Double-barrelled questions (questions that look like a single question but are actually two):

Should the EU cut spending for planetary science and increase spending for health care?

Do you like watching TV documentaries and attending lectures about planetary science?
(Someone might like watching TV documentaries but have no interest in attending lectures.)

Iceberg questions (questions that look simple but are not as clear as they seem)

Do you think space exploration should focus on exoplanets? (Question assumes respondents know what exoplanets are.)

Is it safe for research to be funded by industry? (Which research? Which industrial companies? What do we mean by 'safe'? For whom?)

Hypothetical questions (questions asking respondents to predict future behaviours – you have no way of knowing whether people will follow through)

Will you buy a telescope to go stargazing after visiting our event?

OK, WHAT DO I DO WITH MY DATA NOW?

You'll need to start by entering your data into a programme for analysis. Excel (or free versions such as Google Sheets) are generally fine, or you can also use a dedicated statistics package such as SPSS if you are familiar with it. Allow one row per respondent, with each question having its own column. For some kinds of questions, you may want to give each response its own column indicating simply whether or not that response was selected - this is especially useful when using multiple-response questions where respondents can select multiple options from a list. (We've provided a detailed example below in our case study).

You may find it easier to work with the data if you 'score' it. For example, 1 = yes, 0 = no; or 1 = strongly disagree, 5 = strongly agree. Then, you can compare the percentage responding in particular ways to each question (e.g. the percentage strongly agreeing, agreeing, disagreeing etc with particular statements; or responding yes/no to questions). Drawing on some of the questions above (and a couple of other multiple response questions relating to visitor experiences, such as their motivations for visiting), let's assume the first respondent had the following answers (scores indicated in brackets where appropriate):

How interested are you in going to another talk about supernovas? 3

How clear or confusing were various parts of today's talk?
Introduction: Clear (4)
Today's speaker was an expert: Agree (4)
Visit motivation - family: Not selected (0)
Visit motivation - interest: Selected (1)

The beginning of your spreadsheet would then look like the following:

Respondent	*Interested in another talk*	*Intro clarity*	*Presenter expert*	*Visit Motivation - Family*	*Visit Motivation - Interest*
1	*3*	*4*	*4*	*0*	*1*

If you have collected some background information (e.g. about gender), you can use this as part of your analyses. For example, you could compare the percentages of males and females agreeing with a particular statement. This could include findings such as: '8 of 10 females but only 3 of 12 males agreed or strongly agreed that the introduction was clear'.

Given the recommended length of the survey (short!) and likely sample sizes (quite small), it is unlikely that the data will meet the criteria for statistical testing [1]. However, with well-constructed questions, responses can still be useful. At the same time, it is important not to overclaim from your data. For instance, if people say they intend to do something, that does not mean they definitely will. In addition, even without leading questions, there is often a positive response bias in the data. People are often inclined to agree with (positive) statements so it is advisable to interpret responses with a grain of salt. In addition, you have no way of knowing (unless all/nearly all of your participants filled in the survey) how representative the responses you collected are of the people who were in the audience overall. Moreover, if your audience is the 'general public' (rather than, say, school students) they are generally unlikely to be representative of the wider population (given that people who usually choose to participate in public events often have an interest in the topic to begin with).

GOT IT! HOW CAN I TAKE THIS FURTHER?

Although paper surveys are most common, post-event surveys can also be created online, for completion either via a website (e.g. through providing a weblink), in person using tablets (if you have access to some that can be used at your event, or you're confident participants will bring their own) or for mobile phones. Most modern survey programmes include a portable version, though you may have to pay extra for the relevant App or processing software. When selecting an online survey programme, do keep in mind whether you'll have access to the internet during the actual data collection, or whether you need an offline option that you can then upload once you get within wifi range. Zapier provides a useful review of free online tools to create forms, and apps to capture survey data.

EVENT SURVEYS: EXAMPLE

With thanks to Dr Mark Fuller, UCL, for this contribution

THE EVENT

The Your Universe festival is a three day event (Thursday-Saturday) with schools invited to morning and afternoon slots on the first two days and open to the general public on the Saturday afternoon. The aim is to use space as a tool to help inspire minds to take an interest in physics.

Figure 1 - Stewart Coulter from West Of London Astronomical Society conducting a lens demonstration to help pupils to understand the working principles of a refractor telescope.

WHAT WE DID

Data collection

We collected our data using paper surveys (see our post-event questions file for a copy of the site map and questionnaire that was provided to Your Universe visitors). When possible, pupils completed the first two questions before the festival, which were designed to assess their baseline level of engagement and what they expected to find. After participating in the festival but before leaving they then completed the rest of the survey, which enquired about which parts of the festival they enjoyed and why, facts they recalled and a gauge of excitement about learning more.

Data collection with the public audiences (Saturday) was more limited. We took a head count every 30 minutes, with the age range roughly gauged using a tally chart. We then asked five groups who seemed to roughly match the age distribution to complete the survey questions (verbally).

In addition, we asked festival volunteers and teachers to submit their feedback after the event via an online questionnaire.

Data analysis

We used an Excel spreadsheet (see post-event feedback for details) to enter our data, with one row for each respondent. The first seven columns correspond to pre-event questions A and B and post-event questions A-E. Responses to open-ended questions (post-event questions B and C) were entered by hand. Columns M-Q were used to count and calculate the percentages of students responding to pre-question A about how excited they were before the festival. Finally, Columns S through AB were used to analyse post-event question A, with one column for each possible response (Cosmic time line, Cosmology and black holes, etc). A formula was used to populate the columns, and then to sum the total number of students reporting each activity as their favourite.

WHAT WE FOUND OUT

To our knowledge, this was the first time that evaluation data had been collected in any sort of systematic way for Your Universe. The constant gentle reminder to fill in the evaluations meant that we had an almost 65% return of evaluations, and this proportion was even higher when just considering primary school children.

Far less successful were the post event evaluations sent to the teachers and volunteers, which received only four replies, three from teachers (33% return) and one from a volunteer (out of ten volunteers). Although these are useful as case studies, it is difficult to assess the impact of making changes to timings and structure of the festival based on such a small number of responses.

At the beginning of the festival, students were clearly excited, with 40%, 35% and 25% reporting 'Very', 'Quite' and 'Average' levels of excitement, respectively (pre-event question A). None responded that their excitement levels were 'not excited' or 'bored'.

Although the questions were not worded in the same way (so comparisons must be made cautiously), if we consider pre-event question A and post-event question E as indicative of excitement, we can compare levels of excitement at the beginning and end of the event. Only four reports show a drop in the level of excitement, whereas 19 showed a rise. The remainder showed no change but it should be noted that 17 were 'very excited' at the beginning and end, so no further increase could be measured.

The most popular events seemed to fall into either interactive demonstrations (such as Telescopes and Cosmology) or sit-down presentations (such as Aurora Borealis and Exo-Planets). Looking at students' responses to the open-ended questions, it appears that a mixture of learning something new and the personality of the speaker were the most inspiring aspects.

REFLECTIONS ON USING THE POST-EVENT SURVEY

In the future I would like to investigate ways of capturing more information digitally, as logging responses to paper surveys was very time-consuming. I have used Plickers at other events, which provides data immediately. The return rate also drops dramatically when people take the surveys away with them or are sent digital copies later.
If I were to change the survey, I would consider using fewer open questions, but I would also want to avoid stifling opinions that I had not considered, or influencing or biasing responses by providing too narrow a range of options. I would also reinstate a question (that was removed due to lack of space on the paper copy) about which keywords pupils DID learn about. Making far more use of the pre-event question, 'What are you expecting to learn about today?' would also allow a better analysis of the overall impression that the festival left.

Many thanks to Ben Gammon for his original suggestions of questions to avoid.

Before the Festival

A) How excited are you about visiting the Your Universe Festival Today? (circle one)

Very Excited Quite Excited (Average) Not Excited Bored

B) What do you think you might learn about today? (Circle as many as you like)

Space Travel (Planets) (The Earth) Plasma (The Sun)

Life elsewhere (Chemistry) (The Death of Stars) (How Telescopes work)

(Galaxies) Light Waves (Rockets) (Space Craft) Radiation

(Supernova) (Black Holes) (How old the Universe is) Biology

Aliens Where we come from? Volcanoes (What scientists do?)

STOP…Answer the rest after the Festival

A) Circle your favourite Activity? **(see Map on reverse for reminder)**

1.- Cosmic time line

2.- Cosmology and Black Holes

3.- Telescopes, UCLO

4.- TWINKLE

5.- Stars HR Diagram

6.- LHC Atlas

7.- Robotic telescope

(8.- The aurora display)

9.- Magic planet

10.- Extra Solar Planets

B) Why was this your favourite?

There was a beautiful picture, it caught my attention. Not only that, the explainer was a kind man.

C) Write about one fact you learnt today that you did not know before

That auroras can be seen fully in Iceland, Finland and sweden.

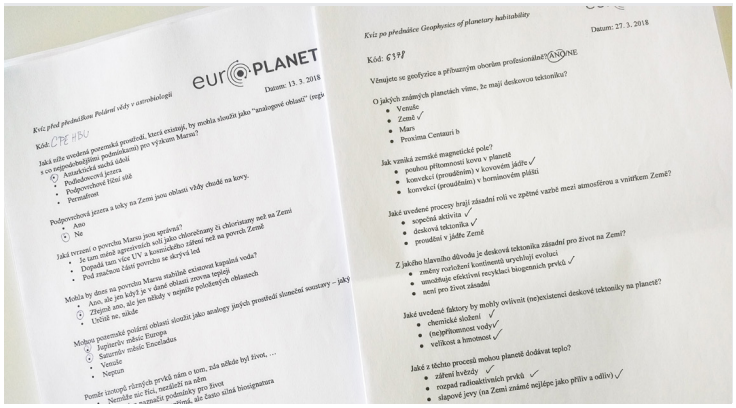
D) Would you like UCL to visit your school to tell you more?

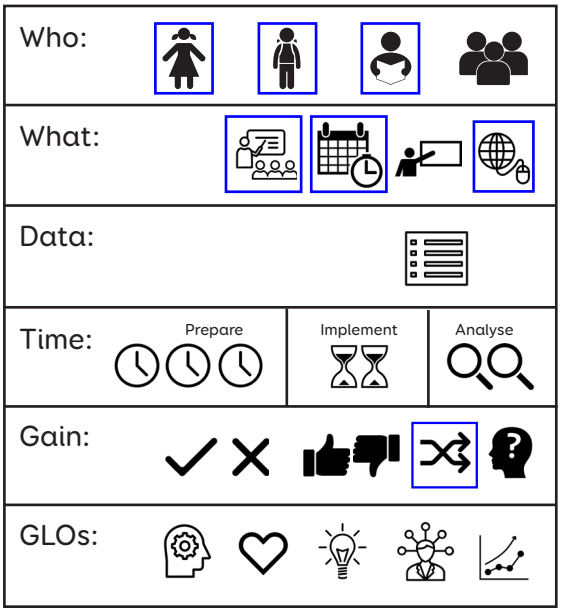
(YES) or NO **(Circle One)**

E) How excited are you to learn more about Space now? (Circle One)

Very Excited (Quite Excited) Average Not Excited Bored

PRE-POST QUIZZES





Pre-post quizzes are brief surveys that are used before and after an event. They are ideal for helping you understand whether or not your audiences have learnt key aspects of the content you are trying to convey. Their purpose is to focus specifically on your content (not on other things your audience might or might not know).

We've included instructions below, and there's also a case study example at the end which uses real data from the Europlanet community to demonstrate how to successfully apply this technique.

WHAT DO I NEED?

Pre-post quizzes can be delivered on paper or online. If you decide to use paper versions, you'll need to prepare plenty of copies in advance and have pencils or pens available at your event, as well as flat surfaces to write on. It's fine if clipboards aren't available as long as there are tables or other surfaces around.

When planning the event, make sure to have time set aside at the beginning and end when people can fill in the quiz!

LET'S GET STARTED

Prior to the event, you need to create the quiz - see our Top Tips section for some essential advice to get you started. The pre- and post-quiz questions need to be the same, so they can be compared. They also need to relate very directly to the content you are covering - this is not the place to find out what else they might know about the topic. For example, if you are giving a talk on exoplanets, you might have three questions related to different aspects, such as the numbers of exoplanets discovered so far, the kinds of stars they have been found near and whether or not scientists have techniques they can use to learn about their atmospheres.

Question types

There are many different question types that could be used, but three of the most common types are multiple choice, True/False and multiple response (see the post-event surveys section for some other question types)

Multiple choice involves a question with different response choices (usually three or four) to select from. There should only be one correct answer. For instance:

What is the name for a region of space that has a gravitational field so intense that no matter or radiation can escape?
a. A green hole
b. A black hole
c. A purple hole
d. A cucumber

A **True/False question** consists of a statement which respondents reply is true or false.

Please select one answer:
The moon is made of cheese. True/False

In a **multiple response question**, individuals can tick more than one response (i.e. there is more than a single right answer). For instance:

Which of the following are planets in our solar system? (Tick all that apply)
Mercury
Venus
Zeus
Earth
Saturn
Aquaman
Jupiter
Neptune

Note: short-answer questions – or any questions where participants have to write in their answers – are not usually suitable for a pre-post quiz as they will take too much time for respondents to complete and it's usually very complicated to accurately compare before and after answers for open questions..

Post-quiz

The questions on your post-quiz need to be identical to those on the pre-quiz, for ease of comparison. The post-quiz can be distributed at the end of the event (either on paper or as an online link participants can access via their smartphones, tablets or computers). For higher response rates be sure people have time to fill them in before leaving the event! You can also choose to run your post-quiz at a later date in order to explore longer term impacts, this is known as a delayed post-quiz.

Comparing responses

Ideally you want to be able to match individuals' responses to the pre-quiz directly with their responses to the post-quiz. However, this means gathering personal information – most likely names – which may be considered unethical (it's better to keep evaluation data anonymous if possible), or may worry the participants as their answers could be traced back to them individually. One alternative is to have participants create a code that is unique to them, but not known by anyone else, that they write at the bottom of their pre- and post-quizzes. For example you could add the following to the start or end of the quiz:

We would like to compare your responses before and after the event. To do this anonymously, please create a code as follows and enter it on the line below.
First 3 letters of your father's first name
First 3 letters of your mother's first name
Last two numbers of the year of your birth
For example, if my father's name is Marco and my mother's name is Sara and I was born in 1983 my code would be MARSAR83.

Or, if your audience is fairly small in size and you think it's unlikely that they'll choose the same code, you can ask them to each come up with their own random suggestion (as Julia did in our case study example below).

OK, WHAT DO I DO WITH MY DATA NOW?

You'll need to start by entering your data into a programme for analysis. A standard spreadsheet package such as Excel is generally fine but you can also use a dedicated statistics programme such as SPSS if you are familiar with it. Allow one row per respondent, with each question having its own column. Score responses to the questions (i.e. '1' for correct and '0' for incorrect) and enter each person's score into the column corresponding to the question. If there are some questions that can be 'proportionally correct', you may need to develop a scale, e.g. from 0-2 (0 if got it wrong, 1 for partially correct, 2 for fully correct; or +1 for each correct response on a multiple response question and -1 for each incorrect response). For example, drawing on the questions above, for a respondent who got the question about black holes correct, the true/false question incorrect and then correctly selected three of the planets but also got one wrong, the table might look like the following:

Respondent	*Black Hole*	*Moon Cheese*	*Planets*
MARSAR83	*1*	*0*	*2*

Then, you can compare the percentage correct (or the scores) for each question, by averaging across all the respondents. If you do have matched data from pre/post quizzes, you can come up with a scoring schema that reflects whether your respondents moved from incorrect to correct (or correct to incorrect) or whether their scores increased or decreased on each question and overall. For instance, a person could receive 1 point per question if they have moved from incorrect to correct - you are basically trying to describe whether their performance improved, decreased or stayed the same. (See our case study example below for further details on how such scoring can be done successfully).

WARNING!! Don't overclaim!

In the likely event that you aren't able to fully match the respondents in your pre- and post-quizzes, you need to be cautious in how you interpret the data. You can't assume that the same sorts of people completed the before- and after- surveys (for example, maybe only people who particularly felt that they learned something took the time to complete the post-event survey?). However, collecting pre- and post-data can still be useful! If you see evidence of improved understanding of your key items across the board, it is likely that at least some of your audience have learnt something - especially if you have good reason to believe that those who responded before and after aren't two entirely different groups of people. (If so, then you need to be really careful!!)

In addition, if your pre- and post-quiz happened fairly closely together in time (especially if the post-quiz was immediately after your event), then you cannot claim a long-term impact. On the other hand, if individuals' performance on the quiz improved from pre to post and the post-quiz was immediately after the event, it is unlikely that the improvement is due to some other cause!

Finally, although the pre-quiz may have drawn people's attention to certain parts of the event (or the content covered), that does not mean that other content was not learnt. Just use your common sense and focus on what data you DO have and what you CAN say about that data when writing up your results...

GOT IT! HOW CAN I TAKE THIS FURTHER?

If you would like, you can incorporate more question types into the mix. For instance, to get a sense of what people may have learnt, and how confident they feel, you could expand on the True/False question type to include the following response choices: Definitely true, Probably true, Not sure, Probably false, Definitely false.

You might also be interested in seeing whether people seem to have retained some of the content learned for a period of time after the event. This is known as a delayed post-quiz. To do so, you could try sending a link to the post-quiz after a period of time has passed (if you have been able to collect email addresses from participants). This would be most suitable for use with an adult audience, though if your audience consists of school children, you can also work with their teacher to distribute the questionnaires. Be warned, however - if you use a delayed post-quiz and see an improvement in responses, this could be because of something people learned later, outside of your event. They could also have looked up the answers when

completing the later survey at home! Finally, when using delayed post-quizzes, it is likely that you will receive far fewer responses - you may need an attractive prize or other incentive to encourage people to submit their delayed post-quiz responses.

If you have data from enough respondents (at least 20 and ideally closer to 50) and would like to delve further into your data, you could utilise some basic statistical analyses. Given the short nature of the quiz, either t-tests or non-parametric versions (Mann-Whitney U test or Wilcoxon signed-rank test) are likely to be most appropriate. Based on feedback from members of the Europlanet community when developing this toolkit, we haven't gone into a lot of detail on statistical techniques here. However there are plenty of good statistical advice resources available; one of the most readable authors is Andy Field.

SIMILAR TECHNIQUES

Quirky questionnaires - Questionnaires conducted in a quirky way related to the topic of the activity, e.g. folded into paper aeroplanes and flown into the 'hanger' (collection box) or 'slam dunked' through a basketball hoop.
Kahoots - Quizzes become more fun with "Kahoots": free editable online games. These are perfect for school students of all ages, and can be run on many different types of devices, making them a very flexible addition to a workshop or classroom environment. For older audiences with their own technology you can also run them using their own mobile phones.

PRE-POST QUIZZES: EXAMPLE

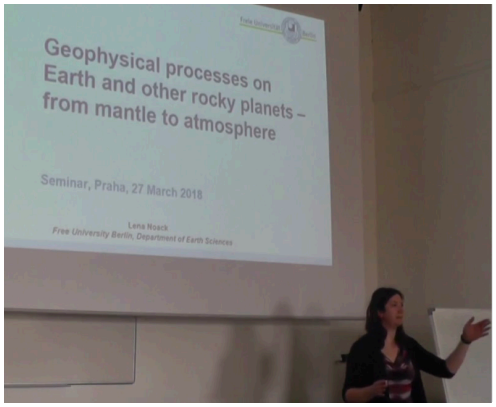
With thanks to Julie Nováková for this contribution

THE EVENTS

Pre-post quizzes were used at three events:

Event 1 (13/3/2018): Astrobiology seminar "Polar sciences in astrobiology" (speaker: Jana Kvíderová)
Total attendance: 12 people (all adults). Quiz responses: 7.
Time: 16:30-18:00
Place: Faculty of Science, Charles University in Prague
Find this talk on YouTube

Event 2 (15/3/2018): Café Nobel talk "The diverse world of exoplanets" (speaker: Julie Nováková)
Total attendance: 35 people (approximately 2/3 adults and 1/3 high school students, a few children with parents). Quiz responses: 15 (some only partial; 8 complete).
Time: 18:00-19:30
Place: Teplice Planetarium



Event 3 (27/3/2018): Astrobiology seminar "Geophysics of planetary habitability" (speaker: Lena Noack)
Total attendance: 25 people (all adults). Quiz responses: 11.
Time: 16:30-18:00
Place: Faculty of Science, Charles University in Prague
Find this talk on YouTube

Further details about the events

The astrobiology seminars are open both to university students (especially attendees of the Introduction to Astrobiology class in the winter semester) and to the general public. Each talk has a Facebook event, and if the speaker agrees, we record the talk and make it public on YouTube. The videos from the previous two years usually have between 300 and 500 views. Julie Nováková is the main organizer, with colleagues Jan Toman and Tomáš Petrásek as co-organizers. The talks take place at the Faculty of Science, Charles University, in late afternoon.
Café Nobel is an activity similar to Science Café and located in northern Bohemia (Ústí nad Labem, Teplice, sometimes other cities). It's organized by a nonprofit of the same name and includes popular talks on both science and humanities (although more frequently science). They're usually attended mostly by adults, partly by older pupils, and sometimes children with their parents.

WHAT WE DID

Data collection

A few minutes before the beginning of each event, I kindly asked the audience whether they would be willing to assist in evaluation of the event by filling a short anonymous pre- and post-questionnaire that would help us see whether they have learned new things there and whether everything in the talk was clear enough or we should improve it the next time. Most people agreed (the response numbers are lower than attendance numbers mostly because some people arrived late or had to leave earlier). I suggested using anything unique as the identifier code - a random number, a favourite character, their initials plus some numbers from their birth date, etc. (I did not suggest only the last option, since using initials and birth date may not seem anonymous to some people, and I was afraid that it would make them think of it more as a test than just a completely voluntary quiz). I emphasized that we're not testing them, just want to improve the event for them. After each talk, I reminded them of the quiz and asked them to please fill in the post-quiz (and leave the
pre-quiz as it was). Afterwards, they handed both papers to me. I waited until the end to collect the papers, since it would delay and disturb the talk at the beginning.

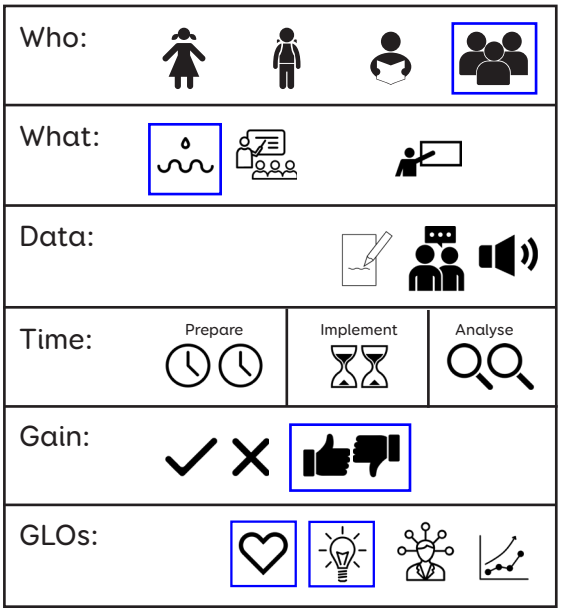
Data analysis

An Excel spreadsheet was created for analysis, with one sheet for each event (we have provided a copy of this spreadsheet at pre-post data). Within the sheet for each event, there is a summary and extra information at the top, followed by scoring criteria ('Quiz overview'). Finally, there is a line for each respondent, with their pre and post scores for each question, which were entered from the paper quizzes. The total points pre-
and post- are found at the end of each row.

WHAT WE FOUND OUT

In all the events, the number of points scored in the post-quiz was generally higher than in the pre-quiz. Sometimes the number stayed the same (especially if the person knew the correct answers already in the pre-quiz), and in only two cases in total, the number dropped (by one and two points, respectively). These drops were related to questions about plate tectonics (twice) and the generation of Earth's magnetic field. Perhaps the discussion of whether Mars could have had plate tectonics in the past or not confused the two people (while the rest answered correctly) and should have been clearer. I am not sure why there was the one wrong response in the other question (saying that Earth's magnetic field is generated by convection in the mantle, while in reality, it's convection in the core). Although there were only 26 complete responses (both pre- and post-quiz fully answered and paired), so any statistical exploration may not be meaningful, I ran a t-test just to see whether the difference between the means (pre: 6.35, post: 8.08) is statistically significant, and it is (two sided $p=0.0024$). Based on that, we can carefully conclude that the attendees indeed learned something new during the talks, and that our approach is most probably a good start. As I mentioned above, we also identified a couple of areas of content that we could perhaps be a little more careful about explaining in future, to try to avoid confusing our audience.

SNAPSHOT INTERVIEWS



Snapshot interviews are very brief, focused interviews, which are used in conjunction with an event to gather impressions quickly, like a photo of a moment in time. They usually consist of no more than three or four questions and should take no longer than 90 seconds to two minutes to complete. They can be used for a range of purposes, such as at the beginning of an event, to gather impressions of why people have come, or at the end, to get an idea of what they thought of the experience. The main thing is to KEEP IT SHORT! This is not the tool to gain deep insight into exactly how people might have learned from an experience, or the long paths that led them to your event.

We've included instructions below, and there's also a case study example at the end which uses real data to demonstrate how to successfully apply this technique.

WHAT DO I NEED?

Snapshot interviews are always conducted in person, at an event. You'll need multiple copies of your questions - enough for one per person you interview. (If you hope to interview 20 people, you'll need 20 copies.) Format the page so that there is space for writing underneath each question (see the case study below for an example of what this can look like).

If you would like to audio record the interviews, you'll need equipment for that as well (modern mobile phones are usually great). However, we have found that writing the answers often helps keep the interviews short!

(Note: While technically these interviews could be done online, the questions are open-ended and thus may appear to require longer answers, which is off-putting to many potential respondents. When interviews are conducted in person, it's more apparent that not much time is required!)

LET'S GET STARTED

Prior to the event, you'll need to prepare your questions. The particular content of the questions depends on what you would like to know. For example, snapshot interviews can be an excellent way of getting to know your audience better. If this is your aim, you might ask questions such as:

Why did you come here today?
Who are you here with?
Have you ever come to something like this before?
What do you think you might find out from this event?

If, on the other hand, you want to gather quick impressions about the event itself, you might ask questions like:

What did you expect this event would be like?
What part of the event did you like the most?
What did you like the least?
Was there anything that surprised you?

Note, although this last question is technically a yes/no question, most people will respond a bit more (e.g. 'Yes, I didn't expect to actually be taken into space.').

It can be challenging to limit yourself to three or four questions. If you find that you would really like to ask several more, it is better to group those questions into sets of three or four and conduct a series of snapshot interviews (i.e. each person could be asked a different set of questions). However, doing so also means that you will be able to gather less data about each question.

When developing snapshot interviews it is also important to note that the questions do not involve further prompts. That is, although it is tempting to ask 'why' someone liked a particular aspect of an event, doing so would take too long and understanding deeper 'why' questions requires different interview techniques (e.g. a longer more in-depth interview). Although snapshot interviews are not suited for understanding such questions, they are good at gathering a range of impressions from larger numbers quite quickly – a snapshot of the landscape of your participants. If your event runs over a long enough period of time, you could perhaps consider doing the snapshot interviews first to get a broader view of your participants' experiences, then following up with more detailed interviews at a later time to probe specific aspects.

For purposes of analysis, it can be helpful to collect a small amount of background information about participants. For instance, if interviewing teachers, it can be useful to add a brief question about which age group they teach (e.g. so you can compare primary with secondary school teachers), or to ask members of a public audience if they have a science degree. You can also make a note based on visual impressions of the age, gender and ethnic background of your respondents, should you so choose. However, snapshot interviews are not the place to ask detailed background questions, for example how long someone has been teaching, in how many schools, when they qualified, how much professional development they engage in and so forth. Likewise, asking several questions to try to pinpoint the socio-economic status or exact nature of the qualifications of members of the public is too intrusive and time-consuming for this type of interview.

At the event (before, after, or both), approach people and ask whether they have 90 seconds (or two minutes) to answer a few questions. Informing them of the very short time frame in advance often means people are more willing to participate. Not everyone – especially teachers – feels they have five minutes to complete an interview, but many feel they can spare one or two minutes! At the same time, you might encounter people who really do want to talk for longer – in that case, it is up to you to decide whether to continue the conversation or whether to try to wrap it up quickly (but politely!) and move on.

OK, WHAT DO I DO WITH MY DATA NOW?

Once you've gathered your data, you need to code it, or group it into categories. Try reading our thematic coding advice to help you with this process. Begin the analysis by reading through your data once or twice, making notes about different categories of responses that seem to be present. Once you feel comfortable with your data, try making a list of categories for it. For instance, these could include things like 'interest', 'fun/enjoyment', 'learning' and 'social'. It is a good idea to create a separate 'master list' of your codes which clearly defines what you mean by each category name and provides an example from your data. (You might find the case study example below helpful in showing how this has worked successfully in the past).

After you have created this set of codes, you need to apply it to your data. As you do that, some categories might be combined, while other new ones might emerge. After a while, you'll probably find that you're working with a consistent set of categories/codes. See our thematic coding advice for further details about coding your data.

Once you have coded your data (put it into categories), then you can summarise it. For example, '8 of 10 respondents said that they came to spend time with their families' or '3 of 15 respondents said the group discussion was their least favourite part of the event.' While it is possible to report these as percentages,

it is not advisable if you have fewer than 100 responses.
You may also be able to look at your data by gender, age etc. However, you do need to be cautious because of how the data has been collected (generally based on visual impressions). Nevertheless, sometimes trends do appear (i.e. if there are clear differences by gender in favourite/least favourite parts of an event).

Finally, although each question is likely to have its own set of response categories, when you look across your analyses, some further connections might be present. For instance, you might find that categories of expectations ('What did you expect this event to be like?') included 'have fun', 'learn something new' and 'using cool stuff' (like looking through a telescope). There could also be categories of what they enjoyed ('What part of the event did you like the most?') that included 'demonstrations' and 'using cool stuff' (looking through telescopes). From this, it is possible to discern that the opportunity to look through a telescope was an important element of the event for your participants, and should probably be continued. You also might want to try to maximise the number of participants who do get to look through a telescope. Note that although it is not advisable to try to combine these kinds of categories statistically, you can discuss such similarities when reporting your evaluation results, especially since they suggest a common recommendation for future similar events.

GOT IT! HOW CAN I TAKE THIS FURTHER?

While snapshot interviews can provide a quick overview of an experience, they cannot provide in-depth insight. However, the responses may provide useful guidance for additional avenues to explore. For example, answers about reasons for attending or who people come with could be used to create response options for self-response surveys. In addition, you may want to understand more about why people responded as they did (e.g. about favourite and least favourite parts of an event) or follow up on other responses (e.g. about their expectations for this kind of event). Such patterns can form the focus of more in-depth qualitative interviews or focus groups[1].

SNAPSHOT INTERVIEWS: EXAMPLE

THE EVENT

This example is taken from a science museum that hosts a lot of visits by school students and their teachers. Their experience is also likely to be similar to a university laboratory or other organisation which is visited by schools.

To improve the experience of the teachers and students who visited - and to be able to encourage even more to attend - we realised it was important to find out more about why teachers were bringing their students and how the visit fits into what they are doing at school.

We also wanted to see if there were differences in timings, especially why some teachers brought their students in the autumn while others preferred the spring/summer, and whether we needed to provide different content at different times of year.

WHAT WE DID

Data collection

Although we wanted to ask a lot of questions, we narrowed it down to four (you can see a copy of the sheet we used to record the questions on the following page):

- What did you do to prepare your students for the visit?
- How is this visit related to the work you are doing in school?
- Have you set your students any tasks to do while they're here?
- Do plan to do anything connected to the visit back in the classroom, after you return to school?

When the teachers first arrived with their students to the schools entrance at our museum we approached as many teachers as we could, and asked them if they thought they might have about two minutes to answer a couple of questions during their students' lunch break. Nearly every teacher we asked agreed - we think this is probably because we emphasised how quick and easy it was going to be. The only exceptions were people who were already planning to leave before lunch or who were not the group's regular teacher (for example they were just accompanying the group as chaperones). We also checked with the teachers

what time they were planning to eat lunch and where we would be able to find them. The museum had a picnic area, which is where most school groups ate lunch.

Later that same day we located the teachers again in the picnic area, and reminded them that they had agreed to talk with us, again emphasising that it would take about '90 seconds, maybe two minutes'. We quickly wrote down their answers (using our snapshots_questions forms) while we chatted and also found out what year group of students they taught. After the interview we thanked them for their time.

Data was collected in the autumn (November) and then again in the late spring (May) so that we could compare the responses from the two different times of year - we wrote the date of each interview on the bottom of each sheet so we could tell the difference between them when we brought it all together.

Data entry

The interview responses were read through and then entered into an Excel spreadsheet for analysis (we've provided a copy of it at snapshots_data). Each interview was entered on a single row, with one column used per question, and a short description of the response to that question typed in (taken from the notes made at the time of each interview). Where the phrasing was verbatim (in exactly the words used by the interviewee) we used 'quotation marks'; other comments were summarised or paraphrased.

We also deliberately left an extra row in between each response, which was used for recording codes, or short labels relating to that response (see 'Data Analysis' below). The data from the autumn interviews were entered on one worksheet (SM-Nov) and the spring interviews on another (SM-May).

Data analysis

Reviewing the responses a few times, we started to see common themes in the responses that the teachers had given. We recorded these on a third worksheet (Codes). We kept re-reading our responses and going back to our codes until we were happy that we had most of the major themes covered within our Codes list. Then we went back through all the data in both worksheets and applied our list of codes to the responses, by writing each code in the row below the responses themselves whenever that theme appeared. Responses could be given more than one code. Finally, we summarised how many responses fell into the various categories. Since we had asked what year group their students were from, and had two data sets from different times of year, we were also able to compare responses between primary and secondary school teachers, as well as between autumn and spring interviews.

WHAT WE FOUND OUT

76 teachers were interviewed (48 primary, 28 secondary). Although we do not know if these teachers are 'representative', we do know that there were more primary school classes visiting than secondary at the times we conducted the snapshot interviews. In addition, the numbers seem sufficient to have captured a range of views, so they meet our purposes quite well.

The main findings included:

- The majority of teachers did at least try to connect the visit to the school curriculum
- Primary teachers (32 of 48) were more likely to mention specific topics than secondary teachers (12 of 28)
- Secondary teachers (10 of 28) were more likely to mention general curriculum connections than primary teachers (9 of 48)
- The unique experience of visiting the science museum was somewhat more important to primary teachers (25 of 48) than to secondary teachers (10 of 28)
- Student enjoyment seemed similarly important to primary and secondary teachers
- Autumn visits seemed to be more academically focused than spring/summer visits, which were more about 'having fun', and with fewer specific tasks set for the students to complete during their visit
- In autumn, 9 of 21 primary teachers and 7 of 13 secondary teachers set students structured tasks to complete during their visit

Moving forward, given that student enjoyment and especially the uniqueness of what is offered by the museum is important to teachers, the museum will try to continue providing enjoyable experiences that are distinct from school. At the same time, we will try to make curriculum connections more explicit - so that both primary and secondary teachers might have an easier time making those links. Doing so is likely to support student learning AND to make it easier for teachers to justify their visit. We also intend to particularly emphasise the worksheets and other academic opportunities we provide during our marketing efforts in the Autumn as these seem of most interest to teachers then.

Science Museum teacher research - snapshot interview questions

Thanks again for agreeing to talk to me, I'll try to make it as quick as possible. We're really interested in why you brought your students to the museum today, and how this visit fits into what you are doing at school. There's no right or wrong answers!

What did you do to prepare your students for the visit?

How is this visit related to the work you are doing in school?

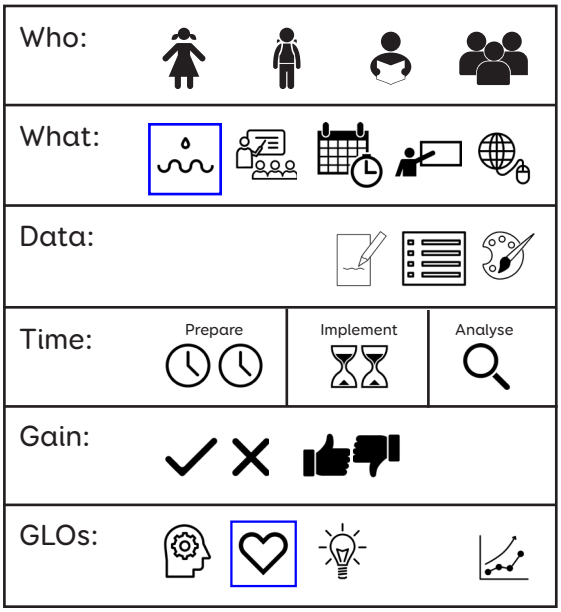
Have you set your students any tasks to do while they're here?

Do plan to do anything connected to the visit back in the classroom, after you return to school?

Finally, what year group did you bring today?

That's all my questions, thanks so much for being involved!
Interview date:

TARGET EVALUATION



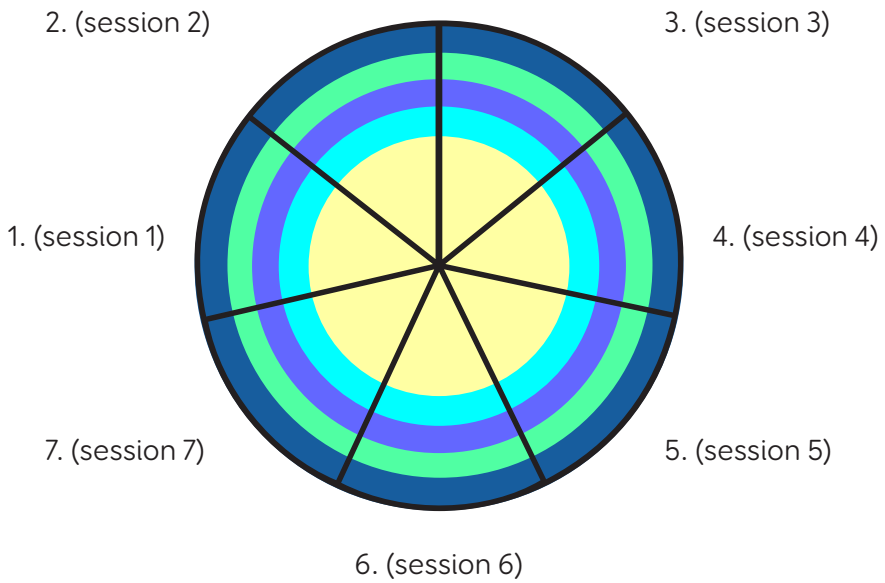
With thanks to Sophie Duncan at the UK National Coordinating Centre for Public Engagement for providing the downloadable example template for this tool.

Visual approaches are often more fun for participants than standard questionnaires, and can quickly allow you to see participants' reactions to your activity. Target evaluation is particularly good for rating different elements of an event, though it can also be used to reflect on overall outcomes or any other questions that can be placed along a rating scale response.

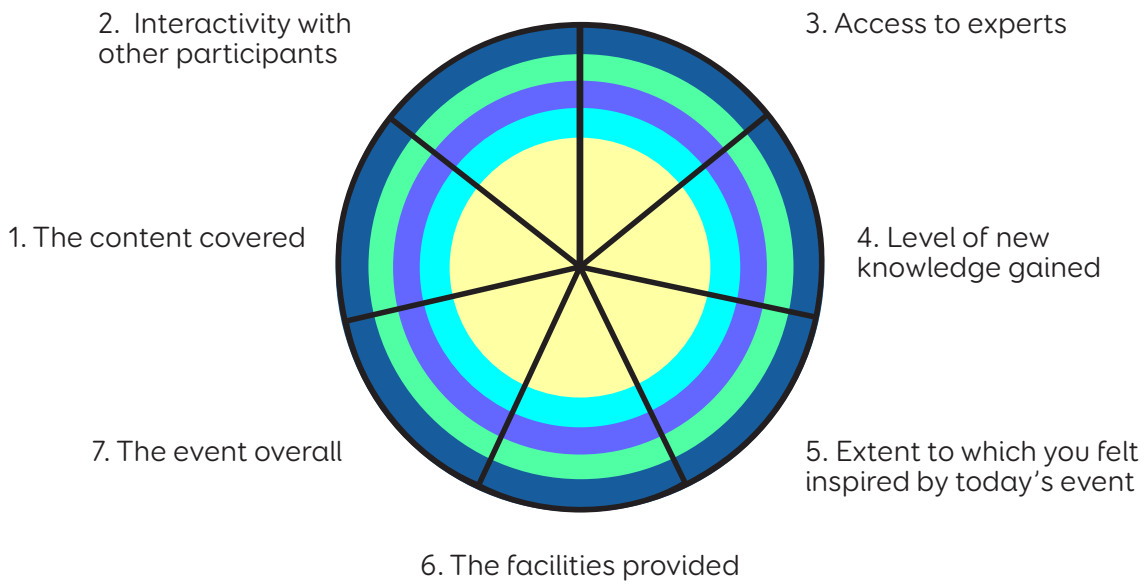
WHAT DO I NEED?

- Pre-prepared large printed dartboards divided into sections (one section per aspect of the event that you're interested in - these could for example be different sessions within an event, or different outcomes or characteristics; see downloadable example for some suggestions)
- Stickers or markers
- Pens

On a scale of 1 to 5 - where 1 is really useful and 5 is not very useful - please rate the following sessions within today's programme



On a scale of 1 to 5 – where 1 is excellent and 5 is poor – please rate the following aspects of the event



LET'S GET STARTED

Distribute your dartboard to your participants – this usually works best in small groups (say 4-6 people each), though can be done individually or as a larger group if necessary.

Ask participants to place one mark or sticker in each segment of the dartboard according to their rating of that activity or outcome, the best being closest to the bullseye, the worst being furthest away. If you print the dartboard out large enough then you can also ask participants to write short notes of explanation near their ranking, to help you understand the reasons behind their score.

OK, WHAT DO I DO WITH MY DATA NOW?

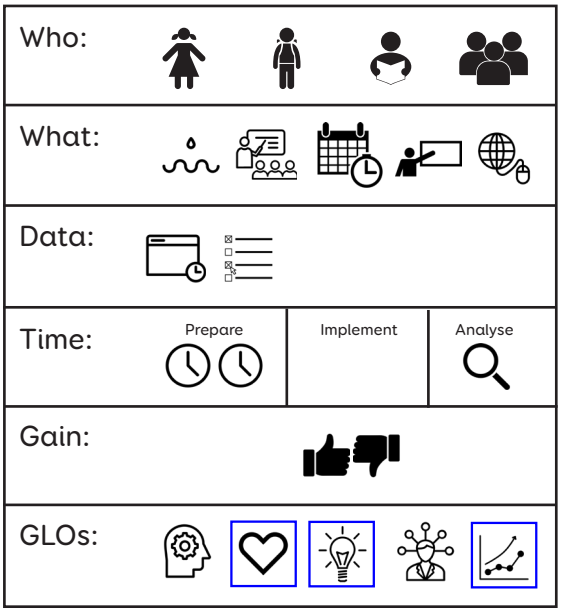
The dartboard approach allows you to very quickly compare ratings for different aspects, allowing you to immediately see particularly strong or weak points within your event, as well as where there were differences of opinion amongst the participants.

SIMILAR TECHNIQUES

Bananas - Give participants a banana (real or paper) and ask them to show them curve up or down to represent happy or sad faces. Participants can then write on the skin any additional comments they might have. If you use real bananas participants can still eat them at the end...
Rate or Slate - This is a great way to get a very visual result on what children think. Children are asked at the end of a session to 'rate or slate' it. Children are presented with a scale labelled 1-5 (1 being negative and 5 being positive.) Children are asked to rate their answer to a question by putting a sticker on the scale.
Facebook Wall - Create a set of "posts" that make evaluation points for the session or event with spaces for dislikes and likes around the post. Then get the young people to like or dislike the post, using markers, stickers or even printed icons in the shape of Facebook reaction options. People can also make comments below each statement using post-its or simply writing on the page.
Picture Voting - Children vote on a range of visual choices around a particular question or issue by attaching stickers to their chosen answer. Choose an issue that can be clearly expressed using pictures of different choices. For example, you might want to find out about ideas for a new lunch-time club or even destinations for a visit. Find as many suitable (and realistic!) images for the children to vote on. Try to find images that are clear and easily recognisable. Find somewhere to display your pictures at the children's eye level where they will all be able to access them. The children can be given one sticker for one vote or you could give them 3 stickers - they could put all 3 on one choice or spread the stickers around if they have more than one 'favourite' answer. Remind the children that there isn't a 'right' or 'wrong' answer.

With thanks to the Woodcraft Folk for the 'Bananas' and 'Facebook Wall' tools and Sheffield City Council for the 'Rate or Slate' and 'Picture voting' techniques.

TWEET SENTIMENT VISUALISATION



Automatic online analysis of tweets to help identify participants' reactions. This analysis must be conducted within a week of the event (whilst the tweets are still "live" on Twitter). It includes a wide range of visualisation options to help classify and categorise the tweet responses, though also relies strongly on choice of an appropriate keyword(s).

Sentiment viz is a free online tool, with further information on its functionality and design also available.

WHAT DO I NEED?

- Access to a computer and the internet
- Identification of an appropriate keyword(s) (e.g. a hashtag used by participants to connect their tweets to your event)

LET'S GET STARTED

Once you've opened sentiment viz you simply type in the relevant keyword(s) and press "Query". The programme then automatically searches the past week of Twitter posts for your keyword(s), and analyses them (see the above links for more details on that process). You then have a variety of different options to choose from in terms of how you want to display your data - see the example at the end of this tool for details.

OK, WHAT DO I DO WITH MY DATA NOW?

Warning!!! This tool isn't perfect. It will collect *any* references to your chosen keyword(s), not just those associated with your specific event or topic. The automated recognition of emotions can also make mistakes at times, but at a broad level it will help you understand how people have reacted.

Warning!!! This analysis *must* be done very soon after the event - you can't back-date the analysis, nor change the time period for which it applies. But with good planning it provides a really simple and easy approach to analysing recent Twitter data to gain broad patterns.

GOT IT! HOW CAN I TAKE THIS FURTHER?

Ideally it would be great to be able to enter *any* qualitative dataset into an analysis tool and automatically produce a meaningful review of key patterns within that data. Unfortunately, at the current time good free online tools for automatically analysing more general text (for example open-response comments from a participant survey) are rare. In general, computers just aren't good enough to accurately identify meaning from short snippets of content. However, efforts are continuing - a more extensive list of current online sentiment analysis tools is also available.

5 — ANALYSIS TECHNIQUES

— Thematic Coding

— Word Clouds

ANALYSIS TECHNIQUE: THEMATIC CODING

A process of categorising open-response (qualitative) data into common themes to assist in identifying patterns within the dataset. Thematic Coding isn't an evaluation technique in itself, but is a useful analysis technique to apply to most sorts of open-response data collected using other techniques (for example, 3 words, Graffiti Wall, Peer Interviews, Photograph Diary and Snapshot Interviews).

Here we take you through the main steps involved - we've also included a case study example which uses hypothetical data from the Europlanet community to demonstrate how to successfully apply this sort of analysis.

WHAT DO I NEED?

- Qualitative responses (preferably in typed form)
- A mechanism to group them together (e.g. highlighters; cut up strips of paper; appropriate computer software)

LET'S GET STARTED

Deeper insights on the success or otherwise of an outreach activity or event can be gained by grouping participants' individual responses to open-ended questions into common categories. This process is the basis of all 'coding' of qualitative data - dividing it into common themes to help better understand the overall patterns in the responses.

The following instructions use a hypothetical dataset of teacher feedback on an imaginary website Europlanet for Schools, as described within the '3 words' and 'word cloud' tools. You might find it helpful to review those tools, and the associated dataset qual coding before continuing here.

First, we need to make sure our data doesn't contain any obvious errors. This could include repeated entries, typos, misspellings or similar. To check our dataset for errors we first need to look at the whole list of words it contains. Within the https://www.wordclouds.com site there is an option to select 'word list' - clicking on this brings up a list of all of the words contained within the data file, as well as their frequencies (how often they occurred within the text). Going back to our data spreadsheet qual coding, take a look at the fourth worksheet, word list, where we have copied the list of words produced by the wordlcouds.com website for our dataset. Looking down Column B we can see that there are a few minor issues with the data, for example:

- There are some small spelling mistakes e.g. relevent / relevant
- Linked but slightly different words have been used e.g. inspirational / inspiring (these are known as "stem words" in language analysis)
- Some phrases / words have the same meaning but have used different words e.g. navigable / easy to navigate / good layout / navigatable

These differences will artificially reduce the frequency of such terms by splitting them over separate entries. Moving on to the fifth worksheet in the qual coding dataset (word list - edited) we have the same list of words, but this time arranged in alphabetical order instead of by word frequency (this helps us to more easily spot all three issues noted above). Column C notes all the places where the original word needs editing, according to the three reasons noted above. The "Updated List" on the same worksheet (columns E and F) shows the same data with all the edits applied - so, for example, 1 count of "accesible" and 3 of "accessible" have been combined into a single entry "accessible" with a frequency of 4.

Let's start the analysis. It is often helpful to begin the analysis by reading through your data once or twice. As you do so, you might want to make a few notes about different categories of responses that seem to be present. For example, if you have asked a question about why people have come to an event, the responses might fall into categories such as interest (responses about interest in the subject matter), fun/enjoyment (thought it sounded fun), learning (thought they would learn about the topic), no choice (someone else insisted they come), social (they wanted something to do with friends or family). In some cases (e.g. if you've

asked what they liked most about an event), responses might occur at different levels. Someone might say they liked a particular aspect best and then say why (even if you haven't asked why). In this case, you might need two sets of categories - one for which aspect they liked the best and one for the reasons. Depending on the way the data was collected, you might not be able to code a response at both levels - that's fine. Just code the data you have. In addition, at times, a response might fall into two categories ('I came to spend time with my daughter, and I wanted to share my interest in Saturn with her'). This might fall under the categories 'family' and 'interest'.

Now it's time for coding... Once you've broadly familiarised yourself with your dataset then you're ready for coding the data into common themes (codes). The sixth worksheet (initial categories) contains the data from the "Updated list" from the previous worksheet (columns A&B). The final step is to go through those data to try to identify common themes. There are lots of ways to do this - some people print them out on paper and use highlighter pens to group connected words, other people cut them up so that each word is on its own piece of paper, and they can be physically grouped together in little piles. In this case we've done it electronically by looking down the "Word" column (B) and thinking about what aspect of the Europlanet for Schools website that response is commenting on. There are no 'right' and 'wrong' categories - the idea is to try to link them in a way that is useful to the purpose of the evaluation. It's clear on reading through the categories that people commented on different parts of the website - for example what it contained (its content), how it looked (design characteristics), and how it worked (its functionality). So, these are three obvious initial categories to start grouping our responses into. Additionally, there are some comments that aren't obviously linked to the website itself, but are more general reactions. In this case we are looking for aspects of the website that participants found worked well, or could be improved, so dividing these reactions into 'positive' and 'negative' is useful for the project team in identifying what to keep, or what to change. Finally, there are a few comments that were either difficult to judge the meaning for (e.g. "cost" or "experience"), or were more general descriptions that didn't fit elsewhere. We've put these under an "other" category. Our initial categories are presented in column C on the initial categories worksheet.

Before going further, it's always worth taking a quick look back at the data to make sure that the categories we've developed are best suited to our purpose. In the case of the qual coding dataset we're trying to get input from the teachers as to what works well, as well as where improvements need to occur within a website. Take a look now at the final revised categories worksheet. First of all, we have taken the list from the previous worksheet and sorted the entries by our initial categories. This allows us to have a better view of each of the words that fit under each category. Looking at 'content' and 'design', most of the words are fairly descriptive. We could perhaps divide some words in the 'content' category into sub-categories relating to 'level' and 'depth' (see column D), but for most purposes this would not be essential. Of more immediate interest however is the 'function' category. By reviewing those words, we can quickly see that there are a range of both positive and negative words - this looks like a clear place where further investigation might be useful, in line with the purposes of this evaluation exercise. Dividing 'function' into 'function - positive' and 'function - negative' and allocating the entries accordingly quickly allows us to see what functional pros and cons the teachers reported during their use of the site.

OK, WHAT DO I DO WITH MY DATA NOW?

Once you have coded your data (put it into categories), then you can summarise it. Note that if you have fewer than 100 responses it's usually best to use the specific numbers in your reporting, for example '15 of 21 respondents felt they had learnt something at the event' or '43 of 48 respondents agreed that they felt inspired by their experience'.

Within the Europlanet for Schools website example, there were 318 responses so we can explore the results in a little more detail. The final revised categories (content; design; function-positive; function-negative; positive-general; negative-general; and other) can be graphed in order to allow us to understand much more clearly how people responded to the website (Figure 1).

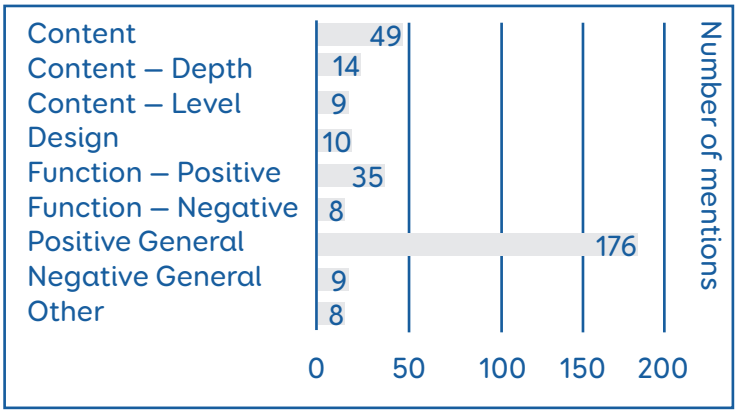


Figure 1 - Teacher feedback to the (hypothetical) Europlanet for Schools website: frequency of final thematic category labels

In particular they help us focus our attention and effort in deciding what actions to take based on the feedback. For example, the overall feedback is extremely positive, which could be very useful for reporting back to a funder to help argue for continued financial support. The content and design of the website in particular appear to be working fine (or at least, any negatives are balanced by positives from other respondents). However, although the functionality is generally well received, it looks like there are some specific elements that could be improved upon - especially the search function, and perhaps some other simplifications. These insights weren't really possible from the more general feedback originally, but help identify specific recommendations moving forwards.

You may also be able to look at patterns in your data according to other demographic categories, for example gender, age or the teacher's subject specialism. Of course this depends on you having collected your original data in such a way as to be able to connect those demographics to your codes! However, you do need to be cautious, especially in the case of methods such as snapshot interviews when the data has been collected in a possibly less accurate way (in the case of snapshot interviews, based on visual impressions). Nevertheless, sometimes trends do appear (i.e. if there are clear differences by gender in favourite/least favourite parts of an event).

GOT IT! HOW CAN I TAKE THIS FURTHER?

Thematic coding is one of the most common (and most useful) qualitative analysis techniques, and there are plenty of further resources available externally if you'd like to extend this approach, for example:

- A great (straightforward and easy to follow) overview of coding open response texts (e.g. from within survey data) is provided at http://www.infosurv.com/how-to-code-open-end-survey-question-responses/.
- For a slightly deeper description of the same process see http://www.getthematic.com/coding-open-ended-questions/.

ANALYSIS TECHNIQUE: WORD CLOUDS

Word clouds are a fun and useful way to easily visualise common themes from data provided by participants in their own words. Though not technically a detailed analysis technique, word clouds do allow you to easily identify data patterns, and are straightforward for anyone to use, making them particularly useful if you haven't used much qualitative data before. Follow the tips and tricks below to make sure you get an accurate result from your word cloud.

WHAT DO I NEED?

- A file containing the responses provided by your participants
- Access to a computer and the internet.

There are plenty of free resources to create your own word clouds available online; two of the best are: https://www.wordclouds.com and http://www.wordle.net

For this example we have used a hypothetical dataset of teacher feedback on an imaginary website Europlanet for Schools, as described in the 3 words tool.

LET'S GET STARTED

Before we start, there are a few important things to know about word clouds:

- There are a lot of different online tools for creating word clouds available, each with slightly different functionality. Some work better with different computer platforms, or different language inputs, or different types of text inputs, so if one doesn't do what you want then try another.
- Note that certain versions limit how much text they actually use in the output they display, or may not be capable of more advanced editing / calculation functions. So it's important to think of them as a "tool" (whose output needs to be checked) rather than a "brain" of their own.
- They typically produce a visual image from a given text input, with the size of each word corresponding to how frequently it appears within the text.
- Most word cloud programmes use "stop words"; these are common words (like "and" and "the") which are automatically removed from the analysis. You can often adjust the stop words if you need to.

OK, here is how to use a word cloud for analysis:

1) If you haven't already, save your data to a file or have it in a format you can easily cut & paste from.
2) Open the word cloud generator you've chosen to use; in this case we'll use https://www.wordclouds.com (as it doesn't require a separate plugin, is usable on most platforms, and contains most of the functions you're likely to need[1]). See the associated dataset qual coding for the example data file, collected using the '3 words' approach. Each worksheet within that file corresponds to a different step in the process as outlined below.
3) Click on 'File' to upload your data (choose whichever option best suits how you have saved the data).
4) Select a word orientation (indicated by the symbol), shape, theme, colours and font to suit your own preferences (in this case we've gone for horizontal writing within a circle shape with a colourful theme in Verdana font). The resulting output displays each word sized according to its frequency; the larger the word the more common it is within our text sample (the colour variation is random). The 'raw data' output from our example data file therefore looks as follows:

[1] If your text is in a different language to English you might want to try wordle.net, as that caters for stop words in many different languages.

Figure 1 - 'Raw Data' output

There are a few issues with this output however... Firstly, some words are repeated (e.g. "Useful" and "useful"), and as highlighted in the top left corner, 3 words were not drawn - the software programme decided they were too large to fit on the screen at the current level of detail. To overcome these problems, we need to adjust our data input a little (this is why it's a good idea to save the data in a file at the start so that you can edit it when necessary...).

Step 1: First of all, we need to shrink the word cloud display to include ALL of the words that it has identified – we do this by adjusting the slider bar towards the top of the page. Keep adjusting it until all the words are included – in this case at a setting of around -35:

Figure 2 - 'Raw data' with shrunken display size to allow all words to be shown

All of the 'missing' words are now included, though the overall size of the cloud is now smaller than in the first image.

Step 2: We still need to fix the repeated words however[2]. Going back to our original qual coding file, on the sheet named Step2 we've applied the function LOWER to convert all the text from the raw data sheet (see the spreadsheet help functions if you're not sure how to do this). Alternatively, you could copy & paste the text into MSWord and then use the 'case' button to toggle all the text to 'lowercase'. (You may need to shrink the word cloud again to include all the words as per step 1)

[2] Some word cloud programmes do this step automatically, or have a function you can apply within the programme itself.

Figure 3 - 'step2' data, with all words now in lower case

Step 3: Looking down the data list in the 'step 2' worksheet shows that some people have entered a short phrase instead of a single word in some entries (e.g. "easy to use"). The word cloud software thinks of this as three separate words ("easy", "to" and "use") so we need to force it to recognise those terms as being connected. The easiest way to do this is to firstly copy and paste the VALUES from the Step2 worksheet into a new worksheet (Step3; use the 'paste special' function in Excel to do this). Then select Edit > Find > Replace and replace all spaces (" ") with a tilde ("~") sign (alt-n on a Mac) - so "easy to use" becomes "easy~to~use". This tells the word cloud software that those terms are now linked, but displays them as spaces for easy reading in the final word cloud (don't forget that you may need to grow/shrink the word cloud display again to get the best view that contains ALL of the words identified):

Figure 4 - Final word cloud, with all words in lower case and phrases recognised using linked words

This is as far as many users will want to go - the resulting word cloud is accurate enough for most purposes, and provides a great visual interpretation of the data from over 100 people.

OK, WHAT DO I DO WITH MY DATA NOW?

The word cloud from the example above was a quick and easy way to understand teachers' common reactions to the website. It is clear that the top two reactions from teachers were that it was 'informative' and 'useful', with many other positive words also being popular ("interesting", "comprehensive", "innovative", "inspiring", "helpful"). This image could be used in a report to a funder or an update to the project team working on the Europlanet for Schools website, to quickly highlight teachers' main reactions.

GOT IT! HOW CAN I TAKE THIS FURTHER?

Further insights on people's reactions can be gained by grouping their individual responses into common categories. This process is the basis of all 'coding' of qualitative data - dividing it into common themes to help better understand the overall patterns in the responses. This approach is described further within the Thematic coding tool.

5 – TOP TIPS

Interested in maximising the success of your outreach evaluation? These tips have been divided into categories according to the various stages of conducting an evaluation, though we do recommend reading them all at the start!

SECTION CONTENTS

5.1 – PLANNING YOUR EVALUATION

- Make sure your evaluation **explores what it is that you're trying to achieve** with your outreach activities, and that it tells you (or others) something useful and/or interesting for planning similar future events. A good starting point is to consider what you want to know and who the evaluation is for (i.e. yourself, funders, managers?)
- **Integrate the evaluation into the activity** as much as possible, rather than have it as a separate "add-on". That way participants feel that it's a natural thing to be part of, rather than something they have to do after the activity is over and they want to leave.
- If you can make the evaluation **easy and practical for participants** to contribute, you will get a better response.
- If possible, try to **use more than one method** for collecting data. Using different techniques will give you different perspectives and better insight into what happened.
- **Only collect the sort of data that you are able to analyse, interpret or display usefully.** It is a good idea to think about how you will analyse the data before collecting it.
- If you are planning to use **online evaluation techniques**, make sure the internet connection/wifi available is reliable and fast enough.
- **Keep it simple!** Don't try to evaluate everything, especially not at first. It is much better to do one thing well - use one method, address a straightforward question - than to try to do too much.

5.2 – COLLECTING YOUR DATA

- **Make it fun** - for both the participants and you! If people are enjoying doing the evaluation they are more likely to take their time and provide meaningful responses. For this reason we've tried to include a number of engaging tools within this Toolkit.
- **Keep the instruments short!** This is not a school exam and if the evaluation takes too long (or even looks too long), people simply won't do it!
- Make sure **the language used in the tools is appropriate** to the reading level of your audience. For instance, for primary school children, sentences need to be kept short and simple. In fact, it's always better to err on the side of simplicity - adult audiences appreciate this too and it makes the evaluation quicker. You may want to consider using a reading age calculator to test how complicated your text is.
- **Pilot your questions and instruments in advance** - it's worth the time it will take. Have three or four people who are similar to your intended audience (e.g. primary school students, members of the general public) try your tool to make sure they understand what is being asked. These people can be family members or friends but should NOT be colleagues or others with similar scientific backgrounds.

5.3 – COMING UP WITH YOUR CONCLUSIONS

- **Don't overclaim from your data!** If people say they intend to do something, that does not mean they definitely will. People are also often inclined to provide positive responses (give you the answer they think you want) so it is advisable to be careful in interpreting some responses, especially if they are about future intentions.
- **Be careful about generalising from your responses**. You usually have no way of knowing whether your respondents were representative of the people who were in the audience overall. Moreover, if your audience is the 'general public' they are unlikely to be representative of the wider population. However, the views they express can still be helpful in understanding their experience! Just be careful to talk about your participants' responses rather than extending your findings to the wider population.
- Analyse your data! Don't simply dump unprocessed data into a Word doc and call it a 'report'. This will make it harder for a reader to comprehend, and may mean that your overall successes get lost. Providing clear indications as to the meaning of your data will ensure that your findings have greater impact.

5.4 – WRITING THE REPORT

- **Include an executive summary** in your report. This will increase the chances that someone (other than you!) will at least have an overview of your evaluation findings. It can also be a good idea to **include different executive summaries for different audiences** (colleagues, funders, managers, even your participants), especially if you're publishing your report online.
- Before starting your report, **ask yourself who do you need to convince and about what?** Then write your report in response to that issue.
- **Use photos and quotes** in your report to convey a real sense of participants' experiences.

5.5 – SHARING YOUR FINDINGS

- Think about how you can **maximise publicity of your findings** within your organisation and beyond it. In addition to circulating the report itself, can you publicise the evaluation via newsletters, blogs and/or journal articles?

RECOMMENDED RESOURCES

Within the Europlanet Evaluation Toolkit we've highlighted key content and practical tools that are most likely to be of interest to the majority of Europlanet members involved in delivering outreach activities. However there are plenty of other good evaluation resources available for anyone who wants to take this further. Listed below are some of the best we have come across.

SECTION CONTENTS

6.1 – GENERAL EVALUATION ADVICE / STRATEGIES

The Center for Advancement of Informal Science Education (USA) produced the Principal Investigator's Guide, designed for **project leaders**. Not assuming any familiarity with evaluation, it provides strategic advice and detailed practical suggestions, including an extensive overview of the Logic Model approach. Also see informalscience.org for what is probably the largest collection of STEM outreach evaluation reports in the world.
The Royal Academy of Engineering (UK) provides a very readable and practical toolkit designed to support **engineers and academics** to evaluate their public engagement activity.
The User-Friendly Handbook for Project Evaluation is an updated version of a classic evaluation guide; this version was commissioned by the National Science Foundation (USA) in 2010.
The European Evaluation Society provides an excellent list of online evaluation books and handbooks including a useful indication as to **who each guide is most suited to**.
The National Co-ordinating Centre for Public Engagement (UK) has compiled a comprehensive summary of evaluation resources suitable for use by **higher education staff**.
RCUK (Research Councils UK) commissioned this guide to public engagement evaluation, explicitly designed for use by **academics and researchers**.
There is a huge range of wider evaluation advice from **outside the public engagement field.** BetterEvaluation is one of the most extensive international resources, comprising 300+ evaluation options and various case studies.

6.2 – EVALUATING ACTIVITIES IN SPECIFIC ENVIRONMENTS

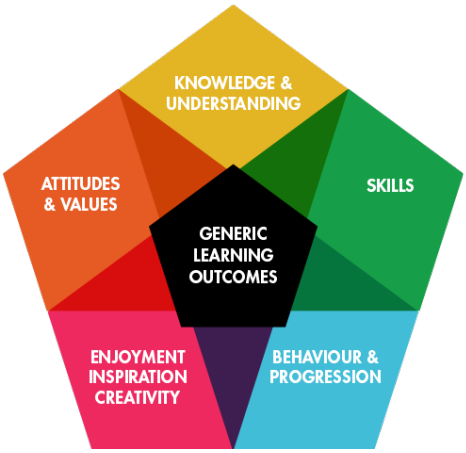
Each of the following resources focuses on one specific activity type or delivery environment.

The EU Directorate-General for Communication has published a toolkit for the evaluation of communication activities (2015). In particular there are separate sections on evaluating **conferences, newsletters, websites, PR and press events, social media and smartphone** activities.

The Cornell Lab of Ornithology (USA) has produced a guide to evaluating Citizen Science projects – whilst you do need to register to download it, it's a very helpful resource if you're at all involved in **citizen science** programmes. It's also worth taking a look at their wider toolkit and resources for examples of evaluation tools that may be suitable to your needs.

6.3 – CLARIFYING YOUR INTENDED OUTCOMES

For your evaluation activity to be both useful and meaningful it is essential that you are clear about what you are trying to achieve: what outcomes (changes to participants) you intend.
One of the clearest and most useful frameworks is the Generic Learning Outcomes, which emphasises the range of emotional, attitudinal, skills and behavioural changes that you might inspire in your participants, in addition to the more obvious knowledge gains:



The above website includes examples of each type of outcome, as well as useful advice on how to apply the framework to your own project, and templates for recording and analysing both quantitative and qualitative data. We have also used this framework in the final columns in our list of tools to explicitly indicate what sorts of tools best suit the main categories of Generic Learning Outcomes.

The EU-funded Space Awareness project is an example of a large-scale (international) outreach programme that applied the Generic Learning Outcomes to create a project-wide evaluation framework. This enabled robust comparison of the project results across a wide range of activity types and audiences, and a clear structure and scope for determining the overall impacts of the project. (The final Space Awareness evaluation report and executive summary are also available publicly).

Warning!!! It is much better to have a small number of prioritised, focused outcomes, rather than an extensive long list - the more outcomes, the more challenging they will be to evaluate!

6.4 — ADDITIONAL SUGGESTED TOOLS AND APPROACHES

There are plenty of other evaluation tools available. Here we link to some of the best technique overviews that we have come across, as well as more detailed advice on specific tools that weren't included in the Europlanet Evaluation Toolkit.

The National Coordinating Centre for Public Engagement (NCCPE, UK) hosts a document published by the Inspiring Learning for All framework which provides a great overview of the strengths and top tips when using a wide variety of different data collection tools.

The European Commission also produced a guide containing 12 separate evaluation tools, including interview, focus group, survey, expert panel, case study, context indicators, SWOT, multi-criteria analysis, and cost-effectiveness analysis.

Additionally, BetterEvaluation has produced a list (and descriptions) of a wide range of other potential tools, themed around where the data is obtained from (individuals, groups, observation, physical measurements or existing records and data).

Online and social media

Though there are other alternatives available, Google Analytics is the classic free **website analysis** software, providing a huge range of information about the online traffic to your site, as well as many useful tools and tips on how to get the most out of your analysis. If your site is fairly large and you're interested in comparing it with other similar sites then SimilarWeb is a great free benchmarking tool providing automated analysis.

If you have a dedicated **Facebook** page then there is a built-in analytics tool called Insights which will give you all sorts of information about how your page content has been used, and by whom. Facebook also provides plenty of support advice on how to access specific Insights information. If you'd prefer a quick overview of the effectiveness of your Facebook communication, including specific recommendations to improve your approach, Likealyzer is a great tool to try, with the added advantage that you can analyse ANY Facebook page (even if you don't hold the admin rights to it!).

Twitter and YouTube also have built-in analytics tools.

Conducting detailed qualitative research

Our initial scoping work suggested that detailed qualitative data collection techniques (such as interviews, focus groups and structured observations) were unlikely to be used by many Europlanet members, so we have not included them in the main toolkit. However, they are an extremely useful approach to gather useful in-depth information, so here are some particularly useful existing resource documents on this topic:

- **Semi-structured interviews**. Design Research Techniques has created a really practical overview of how to prepare and conduct flexible, informative interviews. Further useful advice on this topic is provided by QualPage and the Open University.
- **Comparison of 'good' and 'bad' interview practice**. These videos provide a useful practical demonstration of conducting interviews, and are relatively short (20 minutes total):
  - — Poor practice
  - — Better practice (though not always perfect!).
- **Conducting focus groups**. Citizens Advice has produced an excellent accessible guide on running small-group discussions.
- **Conducting observations**. Taylor-Powell & Steele have created a very practical guide to conducting

observations, incorporating a wide range of logistical and conceptual issues, as well as including some example observation guides at the end.

- **Personal meaning maps (PMMs)**. A very useful qualitative research tool, originally developed for use within museums but now applied extensively elsewhere. PMMs involve two open-ended interviews, conducted before and after an activity, with the participant drawing or writing their responses to a given prompt (key word). Suh (2010) provides a useful overview of how to apply PMMs in practice.
- Our section on thematic coding also includes some great links to other useful resources.

Document authors: Karen Bultitude and Jennifer DeWitt, University College London [May 2018]
This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 654208.