EVALUATION FRAMEWORK & TOOLKIT

IAU OFFICE OF ASTRONOMY FOR EDUCATION
DR. SOPHIE BARTLETT

Contents

INTRODUCTION

Background and Purpose

The aim of this document is to provide a framework and toolkit for NAECs and astronomy education practitioners to aid them in systematically evaluating their education programmes, and provide transferable, evidence-based practices. We adopt a realist approach to evaluation that enables us to account for (and measure) the complexities and subtleties presented across different settings, such as the influence of cultural and practical components that will vary not only across countries, but within countries: across local geographies, schools, etc.

With the evaluation data that is collected, we can work towards measuring the objectives, processes, and outcomes of NAEC activities, and map the cultural and practical boundaries of astronomy education across the OAE network.

We will build an evaluation data repository where NAECs can upload their (anonymised) evaluation data yielded from their astronomy education programmes. Building this evidence/database will facilitate the demonstration of the overarching impact of the OAE on a global level, as well as permit us to perform cross-country comparisons. We intend to use this information to identify what works (and what doesn't), for whom, under what circumstances. Thus, recognising the key criteria for success.

This information will also support the sustainability and legacy of OAE activities. The ability to systematically demonstrate outcomes and impact of initiatives (particularly pilot ones), will add value to applications for further funding.

What is evaluation?

Evaluation has multiple definitions that vary across settings, cultures, and countries. In the context of the OAE, we regard evaluation as a process that uncovers how well something works and the extent to which the intended outcomes of a programme were achieved. Evaluation captures the value, efficacy, effectiveness, and impact of something, such as a teacher training event, or a classroom intervention.

Realist Evaluation

Evaluation can be carried out in various ways with various theoretical underpinnings. We will be taking a 'realist approach' to evaluation across the OAE. When we consider evaluation, the main question we are trying to answer is whether something works as intended, however, the foundation of realist evaluation is that it accounts for particulars, it describes the context surrounding the intervention such as cultural, socioeconomic, and environmental factors. Evaluation therefore seeks to provide *transferable* information, rather than *generalisable* information. Instead of saying 'this intervention works everywhere', we say 'this intervention worked under these circumstances where these processes took place'.

Pawson and Tilley describe realist evaluation as:

> ***What works, for whom, under what circumstances?*** [1]

Realist evaluation comprises of context-mechanism-outcome configurations. The most common component, or one that we are most familiar with, is the outcome. The outcome is essentially the minimum information that is reported from an evaluation, e.g., pupils' interest in astronomy increased. However, an outcome in isolation does

not reveal much about the experiences of the audience, why this outcome was observed, and what the wider implications of such an outcome might be.

For example, we know that science is taught across the globe, but we also know that not all pupils perform equally, nor do they have identical perceptions of science. This is because there are all kinds of other factors at play, not just how science is taught, but time dedicated to teaching and learning science, access to resources, career opportunities, social norms, relevant role models, and many more. Realist evaluation distinguishes these other factors to better understand the ecosystem of influences surrounding your project.

Context-Mechanism-Outcome

We know that an outcome is essentially everything that is yielded as a result of your intervention or activity. Outcomes can be intended or unintended, expected or unexpected. But what is a context and what is a mechanism? Figure 1 summarises the key features of these components.

FIGURE 1: CONTEXT-MECHANISM-OUTCOME CONFIGURATION

Context	Mechanisms	Outcomes
Things that exist outside of your intervention.	The underlying **processes** that take place (within the contextual setting) that lead to particular outcomes.	Everything and anything that happens as a **result** of your intervention.
Context influences whether or not your outcomes are achieved.	Mechanisms are how your audience engages with and responses to your intervention.	Context + Mechanisms lead to Outcomes.
It can be the context that prevents certain outcomes from happening – not solely your intervention.	Mechanisms require the right context, or they will not happen.	Outcomes can be intended or unintended, expected, or unexpected.
You must consider the surrounding context in advance of your evaluation.		

The **context** describes the elements that are separate from your programme and outside of your control but may influence whether or not your outcomes are achieved. It is important to recognise and highlight the context as it may be the context, not your intervention, that prevents certain outcomes from happening. Or it may be that some outcomes will only be achieved through your programme under the right context. [2,3]

For example, your intervention on developing astronomy observation techniques might be very effective for a class that had beautiful clear skies during the intervention, but not very effective for those who had overcast skies. This is a very obvious example, and an outcome that is to be expected. Other situations may not be so obvious. It is important to recognise these different contexts, so you don't falsely draw conclusions about the efficacy of your intervention. It is also helpful to recognise what contexts might not be suitable for replicating your programme.

Mechanisms are the underlying processes (that take place within the contextual setting) that lead to outcomes. Mechanisms describe how your audience engages with and responds to your programme. Mechanisms require the right context, or they will not happen. [1,4]

What does realist evaluation tell us?

Using the context-mechanism-outcome configurations, realist evaluation indicates the conditions under which a programme works, and how it works, as well as the conditions under which it does not work [1]. It therefore allows us to predict whether programmes that prove successful in one setting, may also be successful in another setting. Realist evaluation allows us to uncover the requirements for success and realise the optimal conditions (contexts) for our astronomy education programmes. This is summarised in Figure 2.

FIGURE 2 – WHAT DOES REALIST EVALUATION TELL US?

In this context (C), these mechanisms (M) fired for these participants, generating these outcomes (O). [5]		I.e. In this circumstances (C), these processes took place (M), participants had these responses to the intervention (M), and these outcomes were generated (O).
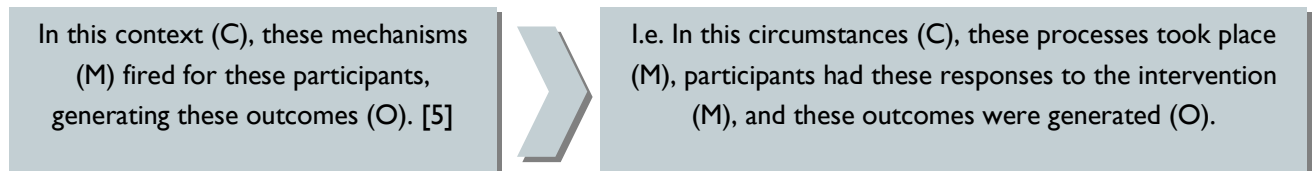
You can view a summary presentation of realist evaluation by following this link.

DESIGNING YOUR EVALUATION

Setting Your Objective(s)

Every intervention and evaluation should start with an objective. A common pitfall in evaluation is that people start designing how they are going to evaluate their intervention, search for surveys or data collection tools they can use, before setting out a specific objective of what they hope to achieve. Your evaluation plan and method should map to your objectives so you can ensure you gather the appropriate evidence.

Practitioners will often have broad, overarching aims such as to promote astronomy teaching in schools, or to promote pupils' interest in astronomy. However, your objectives should be more specific and focused. To do this, we use 'SMART' objectives. SMART is an acronym for Specific, Measurable, Achievable, Relevant and Time-limited objectives. Figure 3 explains these five components that should be explicit in your objectives.

FIGURE 3 – SMART OBJECTIVES EXPLAINED

Specific: What exactly are you trying to achieve?

Measurable: Are you able to measure your progress towards achieving this objective? If you are hoping to induce a change, consider how you will measure this.

Achievable: Is the objective realistic with the time and resources available? Don't aim too high!

Relevant: Are they relevant to your wider ambitions and to your audience? If you achieve your objective, it should be of benefit to your audience.

Time-Limited: When is the deadline for your objective?

An example of an aim or vision of your NAEC team might be to increase teachers' confidence in teaching astronomy. One project contributing towards achieving this aim involves running a one-day teacher training event focused on the Earth-Moon-Sun system and seasons on Earth. You will therefore want to set out a SMART objective for this one-day event. Figure 4 provides an example of what your SMART objective might be for this particular event.

FIGURE 4 – SMART OBJECTIVES EXAMPLE

Specific: Increase TTP teachers' confidence in teching the Earth-Moon-Sun system

Measurable: Implement a survey before and after the TTP event that measures confidence

Achievable: The TTP focuses on the Earth-Moon-Sun system and methods of teaching these concepts

Relevant: If teachers feel confident in their delivery of particular teaching methods, then they are more likely to implement these methods and share them with colleagues

Time-Limited: Teachers' confidence will have increased by the end of the TTP in November 2023

Creating a Logic Model

Now that you have a clear objective for your programme, you can begin to construct your evaluation plan. A popular and useful approach is to create a logic model. Logic models are a graphic illustration of the fundamental circumstances, resources, activities, and intended outputs, outcomes and impacts surrounding a particular programme. A template is provided in Table 1. Logic models are not only helpful in the evaluation of a project but also in the initial design and planning, as well as its implementation. They help to ensure that evaluation is embedded in your programme and implementation plan, and not just an after-thought or add on.

TABLE 1 – TEMPLATE LOGIC MODEL

Your planned programme		Your intended results		
CONTEXT The context and circumstances surrounding your programme. The problem you are trying to address or gap you are aiming to fill. **ASSUMPTIONS** Facts or conditions that you assume to be trust and that will facilitate the success of your programme.	**RESOURCES/INPUT** Existing resources that can be used in your programme. e.g., funding, staff time, venue space, expertise, etc. **ACTIVITIES** What the activity / event / initiative will involve. Actions carried out from day one of the programme.	**OUTPUTS** Anything generated as a direct result of your programme. e.g., a project, website, resource, etc. Should be directly observable.	**OUTCOMES** The results of changes that occur as a direct result of your activity. Will typically relate to your audience or participants. You should be able to measure or observe outcomes. Outcomes should follow your 'SMART' objectives.	**IMPACT** Broader impacts and inferences as a result of your outcomes. You may have short- medium- and long-term impacts.

Logic models are well suited to realist evaluation. They facilitate the process of defining your intended outputs and outcomes, but also ensure that you consider and identify potential factors surrounding your programme that might influence your outcomes.

You should create your logic model at the earliest opportunity in your programme. The context, assumptions, and resources/input components of your logic model are the 'knowns' surrounding your programme. The outputs, outcomes and impacts are what you expect or predict to achieve through the programme's implementation. When considering your outcomes, keep in mind that it is important you are able to observe or measure these outcomes and thus demonstrate whether they have been achieved. The outcomes should be based on your SMART objectives.

It can be easy to confuse outcomes with impacts. Consider the impact of your project as relating to the broader aims and vision of your NAEC team or organisation. These are often broader targets you are working towards and may result from your outcomes. You will likely have other projects and activities that are also contributing to these intended impacts.

As an example, you might be running an astronomy camp for children aged 9-14. Table 2 provides an example of what your logic model might look like.

TABLE 2 – EXAMPLE LOGIC MODEL

Your planned programme		Your intended results		
CONTEXT Low proportion of children pursue STEM subjects. Children lack understanding of the applications of astronomy. Many children lack access to dark skies. **ASSUMPTIONS** Parents desire summer activities for their children. Astronomy is interesting and exciting to children. Favourable weather will permit observing nights.	**RESOURCES/INPUT** Funding, e.g. from OAE. Experience of delivery team in similar activities. Access to equipment, e.g. telescopes. Volunteer network / amateur astronomers. **ACTIVITIES** A summer astronomy camp for children. Multi-disciplinary, hands-on, creative activities. Observing nights.	**OUTPUTS** Resources / activities. An astronomy camp 'how to' / curriculum (for repeat implementation).	**OUTCOMES** Children enjoy the camp and are enthused by astronomy. Children make friends, feel a sense of belonging. Children gain / increase exposure to dark skies. Children have increased understanding and interest in astronomy. Children have increased understanding of the applications of astronomy.	**IMPACT** Short-term Children and family seek out further astronomy and STEM activities. Medium-term Increased pupil engagement with STEM lessons at school. Long-term Increased uptake of astronomy and STEM courses

Data Collection: Quantitative or Qualitative

Data is either quantitative or qualitative. Evaluators may collect one or the other, or a combination of the two. Whether you collect quantitative or qualitative data will depend on your objectives and what information you want to obtain from your audience. In realist evaluation, practitioners often use both quantitative and qualitative methods [5]. Contexts and outcomes are often yielded through quantitative data, and mechanisms, i.e. the processes that took place during your programme, are best captured through qualitative data. Qualitative methods are crucial if one of your objectives is to understand how or why something worked or did not work.

Quantitative methods are suited to studies that seek to test hypotheses and measure numerical change. Typical tools for quantitative data collection include tests or quizzes, closed questionnaires, and numerical observations (for example, counting how many times pupils ask a question during a lesson, or counting how many times pupils interact with their peers). Table 3 summarises some key advantages and disadvantages of quantitative data collection methods which may be helpful when deciding if they are suited to your project.

TABLE 3 – ADVANTAGES AND DISADVANTAGES OF QUANTITATIVE METHODS

Quantitative Methods	
Advantages	**Disadvantages**
Can provide an overview of your full sample.	Rigid approach – participants cannot explain/justify responses and can only provide responses to the particular questions asked.
Statistical analysis is considered rational and scientifically objective.	Data collection is not natural – context is structured.
Quick process – you can collect responses to 100 surveys simultaneously.	It is difficult to design good quantitative data collection instruments.
Easy to replicate.	Requires a larger number of participants.

Qualitative methods involve non-numerical data. Common examples include interviews, focus groups and observations (e.g. observing interactions among your participants both with each other or with your activity). Qualitative data is often presented as text and/or documents. Even though interviews are conversational, typically they will be recorded and transcribed. Qualitative methods permit more expressive data collection and therefore assign more freedom and autonomy to your participants in terms of what information they share. Table 4 summarises the advantages and disadvantages of qualitative data collection methods which may be helpful when deciding if they are suited to your project.

TABLE 4 – ADVANTAGES AND DISADVANTAGES OF QUALITATIVE METHODS

Qualitative Methods	
Advantages	**Disadvantages**
Requires less participants than quantitative methods.	Time consuming / expensive.
Yields understanding of why things occurred and underlying processes.	Subject to bias – greater difficulty in proving reliability and validity.
Permit exposure to subtleties and complexities.	Hard to gain representation of a group or sample.
Can be flexible and changed.	Can be difficult to draw generalisable conclusions.

Who can Provide Evaluation Data?

It is also important to consider who might be best placed to provide you with the meaningful data that tie to the objectives of your programme. There are generally four key participant groups to choose from:

Your **immediate audience**, those who are directly engaged with or receiving your intervention or project.

The **family members** of your audience. This is a particularly helpful group when you are engaging with young people as they can provide valuable information about your audience.

The **key stakeholders** surrounding your project. These are individuals who have some involvement or a 'stake' in your project. E.g., sponsors/funding bodies, teaching staff, venue staff, etc.

You as the education practitioner / facilitator can also provide valuable evaluation data about the delivery and success of the project by providing detailed reflective accounts.

When deciding who to gather evaluation data from, it is helpful to consider the following:

- **Your objectives and intended outcomes**: who can provide the appropriate information to reveal whether these objectives have been met and the outcomes achieved?
- **The accessibility of your audience**: is there a method of data collection that is both suited to the data you require and to the participants?
- **The nature of your activity and audience participation**: is there opportunity during your project to reach these individuals and gather data? How will you embed your data collection into your project and audience activities? The more straightforward, the better.

When to Collect Evaluation Data

It is important to consider when, in the context of your programme, you will collect your evaluation data. Depending on the nature of the activity/intervention you are implementing. There are distinct timepoints to consider, these are before (pre) your programme has started, during your programme, immediately after (post), or a while after your programme has finished (follow-up). These are summarised in Table 5. Each timepoint will provide different information from your participants so it is important to consider what information you need to demonstrate your outcomes.

TABLE 5 – DATA COLLECTION TIME POINTS

Timepoint	Description	Information you can gather
Pre	Pre refers to pre-event or pre-programme. This is the data you collect before your audience has engaged with any of your programme activities. Pre data gives you a baseline understanding of your audience.	• The baseline or benchmark of your audience's perceptions, attitudes, or understanding. • What your audience hopes to gain from your event. • Your audience's expectations of your event.
During	This is the data you collect while your programme or event is ongoing. The audience provides this data whilst they are engaging in your programme.	• What your audience is doing. • Your audience's current experiences of your programme.

		• How your audience is engaging with your programme.
Post	This is the data you collect after your audience has finished engaging with your programme or activity. Post data gives you an understanding of your audience following their experience.	• Your audience's perceptions, attitudes, or understanding following your programme. • Your audience's experience of your programme.
Pre and Post	Repeating data collection methods both before (pre) and after (post) your audience has engaged with your programme can help you identify if their views or understanding has changed. You can compare data from the two time points to see if there are any differences.	• Collect the same data (ask the same questions) pre- and post-programme to make direct comparisons. • Whether, and the extent to which, your programme had an impact on your audience's perceptions, attitudes, or understanding.
Follow-Up	Collecting data after time has passed since the end of your programme is called follow-up data. It is relatively easy to have a short-term impact on your audience and have them leave your programme or event excited, enthused or with new knowledge, however, you may want to find out if that impact is long lasting and/or if they remember anything further on, perhaps a few months or a year later. Where you can do a comparison between pre and post data, you may also want to do a comparison between post and follow-up data.	• What your audience remembers from your programme or event e.g. six months later. • If you collected post data (immediately after your programme), does your audience still report the same perceptions, attitudes, or understanding after several months? • Have your audience changed their actions or behaviour in anyway as a result of participating in your programme?

The time at which you plan to collect data from your audience and the objectives of your data collection will influence what kind of questions you ask, and the information you want to gather.

For example, if you are planning to only collect data after (post) your programme, and not before, then you will need to consider how you will identify what, if any, influence your project has had on your audience.

Let's assume one of your intended outcomes is to increase participants' enjoyment of astronomy. If you are collecting evaluation data both before and after the intervention, then you might ask question A in Table 6 on both occasions. This will allow you to compare participants' responses on the two occasions, and hopefully they will have a greater level of agreement after your intervention.

If you are only collecting data at the end (post) of your event, then you will likely need to amend the wording of your question so that it asks the participant to reflect on how they feel the intervention has influenced them. Here you may want to use an example similar to question B in Table 6.

TABLE 6 – DIFFERENT QUESTIONS FOR DIFFERENT TIME POINTS

	Question / Statement	Response Options
A	I enjoy astronomy	Strongly disagree Disagree Neither disagree nor agree Agree Strongly Agree
B	Having done the activity, I enjoy astronomy…	Much less Less Neither less or more More Much more

Ethical Considerations

A vital thing to consider when conducting evaluation and gathering data from audiences is whether you need to follow specific ethical protocols. The OAE has obtained ethical approval from Cardiff University to use the data collection methods in this toolkit and use the anonymised data you collect for analysis and publications. Publications include any material that is shared and viewable outside of the OAE, such as reports, presentations, and journal articles.

> **If you intend to use this toolkit to evaluate your programmes, gather feedback from your audiences, and publicise your findings, you must follow the ethical guidelines we prescribe, and complete a NAEC Evaluation Agreement form.**

All information on the ethical protocols is stored in this Google Drive folder. Sophie Bartlett will also be running online webinars on the ethical protocols, so please contact her if you plan to use this toolkit and publish your results.

If you collect data without following these ethical protocols, your data will not be eligible for publication.

Evaluation Toolkit

The remainder of this document includes the toolkit of data collection instruments for gathering information and feedback from your audiences. What data collection instrument you use will depend on a number of variables, such as your objectives, your audience, the format of your event, and the time and resource you have available. The next two pages provides a tool selector matrix. Along the top row of the matrix (table), you will see the eight data collection tools prescribed in this toolkit. The left side column displays various characteristics and formats of education programmes. The matrix will help you to identify which data collection tools are suitable for which education programmes.

After the matrix, each of the eight data collection tools are described in terms of how they can be implemented, what equipment/material you will need, some examples and suggestions for questions/prompts, how to analyse your data once you have it, and some advantages and disadvantages of each tool.

Tool Selector Matrix

		Graffiti Walls	Mentimeter	Surveys	Interviews	Peer Interviews	Journal / Diary Entries	Tests / Concept Inventories	Drawing
Audience Group	**Primary / Elementary School Pupils**	✓	✓	✓	✓	✓	✓	✓	✓
	Secondary / High School Pupils	✓	✓	✓	✓	✓	✓	✓	✓
	College or University Students (age 16+)	✓	✓	✓	✓	✓	✓	✓	
	Teachers	✓	✓	✓	✓		✓	✓	
	Teacher Trainers	✓	✓	✓	✓			✓	
	Families / Communities	✓	✓	✓	✓	✓		✓	✓
Event type	**Classroom Workshop**	✓	✓	✓	✓	✓	✓	✓	✓
	Teacher Training	✓	✓	✓	✓		✓	✓	
	Lectures / Talks / Shows		✓	✓	✓			✓	✓
	Webinars	✓	✓	✓				✓	
	Website / Blog			✓				✓	
	Competitions / Contests			✓	✓	✓		✓	
	Internships / Placements			✓	✓		✓	✓	
	Social Media Campaign / Activity			✓				✓	
	Festivals / Celebration Events	✓	✓	✓	✓	✓		✓	✓
	Astronomy Camps	✓		✓	✓	✓	✓	✓	✓
	Online Videos			✓				✓	

		Graffiti Walls	Mentimeter	Surveys	Interviews	Peer Interviews	Journal / Diary Entries	Tests / Concept Inventories	Drawing
Objective (pupils)	Increase Understanding of Astronomy	✓	✓		✓	✓	✓	✓	✓
	Increase Practical Skills in Astronomy				✓	✓		✓	
	Increase Interest	✓	✓	✓	✓	✓	✓		✓
	Increase Confidence	✓	✓	✓	✓	✓	✓		
	Increase Future Aspirations	✓	✓	✓	✓	✓	✓		✓
Objective (teachers)	Increase Understanding of Astronomy	✓	✓		✓		✓	✓	
	Increase Application of in Astronomy to other Subjects	✓	✓	✓	✓			✓	
	Increase Confidence in Teaching	✓	✓	✓	✓		✓		
	Improve Teaching Practices				✓		✓	✓	
	Increase How Much Astronomy is Taught in Schools		✓	✓	✓		✓		
When	Before		✓	✓	✓	✓	✓	✓	✓
	During	✓	✓	✓	✓	✓	✓	✓	✓
	After		✓	✓	✓	✓	✓	✓	✓
Time Commitment	Preparation	🕐	🕐🕐	🕐🕐	🕐🕐🕐	🕐🕐🕐	🕐	🕐🕐	🕐
	Implementation	🕐	🕐	🕐	🕐🕐🕐	🕐🕐🕐	🕐	🕐	🕐🕐
	Analysis	🕐🕐	🕐	🕐🕐	🕐🕐🕐	🕐🕐🕐	🕐🕐🕐	🕐🕐	🕐🕐🕐

GRAFFITI WALLS

DESCRIPTION

Graffiti walls involve using large sheets of poster paper or post-it notes that the audience uses to share their experiences of an event. They are particularly suited to one-off events with large audiences, such as science festivals or online public events.

WHAT YOU WILL NEED

In person events:

- Large wall space
- Post-it notes / poster paper
- Pens
- Questions to prompt particular responses

Online events:

- Access to a relevant website, e.g., Padlet or Miro (these have free versions)

IMPLEMENTATION

For in-person events, you should have at least one dedicated space at your venue for people to contribute to the graffiti wall. You may want to have several graffiti walls at your venue to capture audience feedback on more than one topic. For an online event, you can use a purpose-built software such as Padlet or Miro.

SUMMARY	
Audience Group	Elementary School Pupils High School Pupils College / University Students Teachers Teacher Trainers Families / Communities
Event Type	Classroom Workshop Teacher Training Festivals / Celebration Events Astronomy Camps
Objective (pupils)	Increase Understanding Increase Interest Increase Confidence Increase Future Aspirations
Objective (teachers)	Increase Confidence
When	During Event
Time	Preparation: Low Implementation: Low Analysis: Medium

You can guide the audience to share information around particular aspects of your event by posing questions or prompts to them. The prompts you use to capture information through your graffiti walls should map directly to the objectives of your event. For example, 'what did you like most about today's event?', or 'what surprised you today?'. Your audience can respond to these questions either by writing on a big sheet of paper, or using post-it notes. If you want to distinguish between groups within your audience, you can use colour codes, for example have male participants use green post-it notes and have female participants use red post-it notes.

EXAMPLES AND SUGGESTIONS

Table 7 provides some examples of prompts or questions you may want to use for your graffiti walls. You may want to adopt these, or they may simply give you some inspiration for more suitable questions for your own programme and your specific objectives.

TABLE 7 – EXAMPLE QUESTIONS AND PROMPTS FOR GRAFFITI WALLS

For young people / family events	For teachers
Finish the sentence: • Today, I liked being able to… • Today, I was surprised to find out that… • Today, I was most impressed by… • Today could have been better if… • Following my experience today, I intend to… • I think astronomy is… Ask specific questions: • What did you like most about today? • What would have made your experience of today better? • What can you tell us about [relevant topic, e.g., the solar system]? • What would you like to be when you grow up? • What did you learn today? • If you had to describe your experience of today in one word, what would it be?	Finish the sentence: • Following my experience today, I intend to… • Today, I most liked… • Today could have been better if… • Astronomy as a context for teaching science is… Ask specific questions: • Was there anything from today that you plan to implement in your classroom? • What is your view on the accessibility of the learning resources that were shared today?

ANALYSIS PROCESS

You can choose how much depth you want to go into when analysing graffiti wall data. A straightforward approach is to log all the key words that participants provide in their responses and make this into a word cloud. Word clouds are a visual representation of all the words included in responses and words that are more frequent in the data appear larger in the cloud. You will find various online software to help you do this.

If you want to explore your responses in more depth, you can look for patterns and commonalities in the responses. You can group responses into categories or themes, for example, have people in your audience responded to your questions in similar ways? What terms or phrases are they using? What overarching themes can you identify from participants' responses?

ADVANTAGES	DISADVANTAGES
• Easy to implement during events. • Interactive and fun for participants.	• Can be hard to distinguish responses between audience groups – e.g., male/female, adult/child. • Requires encouragement to ensure audience engagement.

MENTIMETER

DESCRIPTION

Mentimeter is a free online presentation tool designed to capture audience responses in fun, interactive ways. Audiences vote or answer questions, and their combined responses are projected back to them and visualised in real time. Responses can be displayed in a variety of formats such as word clouds and bar charts.

Mentimeter is particularly useful for online events but can also be implemented at in-person events if you have access to a laptop and projector.

The tool is good for summary, snapshot data collection rather than specific, in-depth data, or data you want to collect over a prolonged period of time. In this latter case, a survey might be more suitable.

WHAT YOU WILL NEED

- A free account with Mentimeter / Kahoot.
- Audience members will need access to a smartphone or tablet.
- If an in-person event, a laptop and projector to display the visual presentations.
- A series of questions to ask your audience.

SUMMARY	
Audience Group	Elementary School Pupils High School Pupils College / University Students Teachers Teacher Trainers Families / Communities
Event Type	Classroom Workshop Teacher Training Lectures / Talks / Shows Webinars Festivals / Celebration Events
Objective (pupils)	Increase Understanding Increase Interest Increase Confidence Increase Future Aspirations
Objective (teachers)	Increase Understanding Increase Confidence
When	Before Event During Event After Event
Time	Preparation: Medium Implementation: Low Analysis: Low

IMPLEMENTATION

Prepare your Mentimeter questions and slides ahead of your event. You can choose from a number of display formats, such as bar charts, scales, and word clouds. During your event, make sure you 'present' your Mentimeter slideshow so everyone can see it (on a projector at an in-person event, or share your screen for an online event). To participate, your audience will need to go to the web address 'menti.com' on their phone, tablet, or PC, and input the code that your Mentimeter slideshow will have automatically generated. They can then respond to your questions and see anonymised versions of their and their peers' answers appear on your Mentimeter slides. Audience responses will save automatically on your Mentimeter account so you can extract them later.

EXAMPLE STATEMENTS AND QUESTIONS

Table 8 provides examples of different styles of questions you might want to use for the different display options on Mentimeter, as well as some example questions and prompts.

TABLE 8 – EXAMPLES OF QUESTIONS AND DISPLAY OPTIONS ON MENTIMETER

Word clouds

These work best when questions encourage one-word responses. Words that have been used more frequently by the audience appear bigger than words used less frequently.

Examples:
- What three words would you use to describe your experience today?
- How did you feel in today's session?
- What three words do you associate with astronomy?
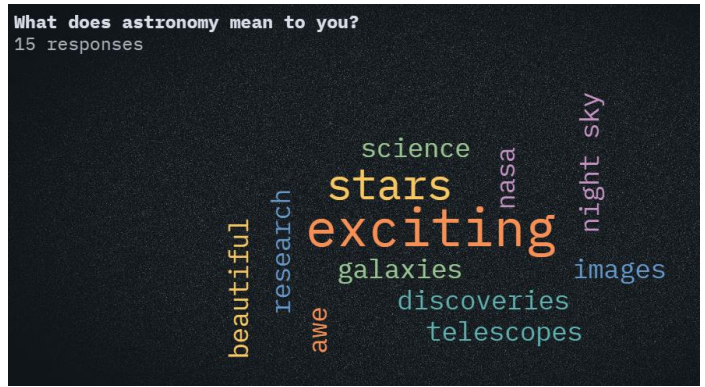
Scale questions

Participants respond to questions or statements on a linear scale and the average scores are displayed on screen. Scale questions require you to specify the question (how confident do you feel carrying out the following tasks?), the components (submitting an observation request, producing a colour image, etc.) and the two ends of your scale (not at all confident, entirely confident).

Examples:
- How confident do you feel in…
- How easy / difficult was…
- How much do you like / dislike…
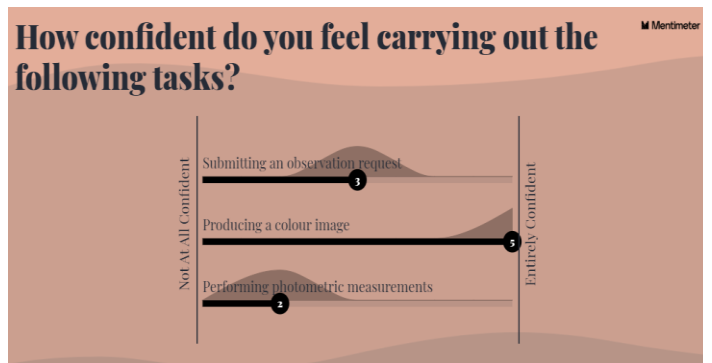
Open text comments

Open text comments give participants more flexibility in their responses as they are not constrained by choosing from pre-populated answers. It also gives participants an opportunity to explain their views.

Examples:
- What did you like most about today's session?
- What could have been better about today's session?
- What, if anything, stuck with you today?
- How would you describe your experience today to a friend?
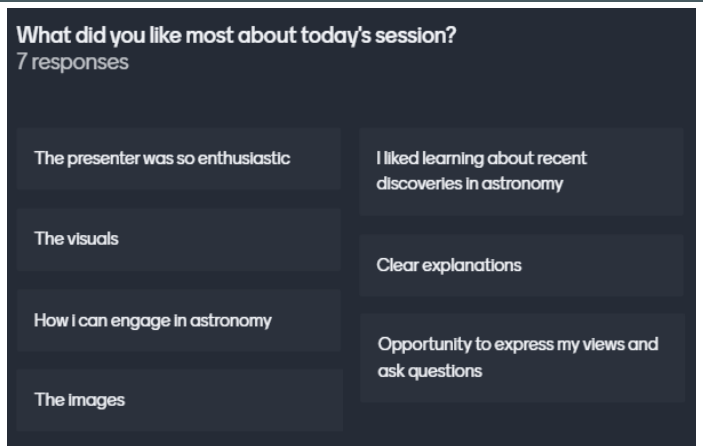
Multiple choice questions

These are closed questions where participants must pick one response from several options. The results to multiple choice questions can be displayed in a variety of formats. The example here shows a bar chart. You will need to specify the question and provide the different response options for participants to choose from.

Examples:
- Which is the most abundant element in the Sun? [hydrogen, helium, oxygen]
- What did you think of today's session? [I didn't like it, I'm not sure, I liked it]
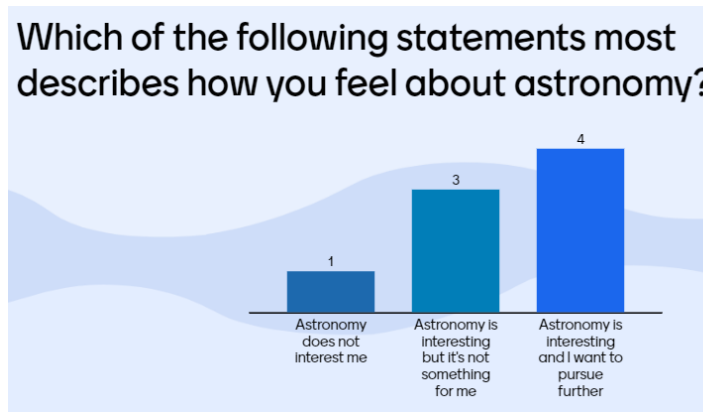
ANALYSIS PROCESS

The great thing about Mentimeter is that it does some analysis for you and, where appropriate, will even present them in appropriate graphs/charts. If you use Mentimeter on several occasions, it might be helpful to compare results. For example, if you implement with different groups, you could compare the groups. If you implement with the same group over multiple occasions, you could explore changes in responses with time.

Mentimeter will save the responses you gather in your Mentimeter account so you can access it whenever you need to. You can download this data either as a PDF or Microsoft Excel file.

ADVANTAGES

- Quick and easy to implement and gather data.

- Interactive and fun for participants.

- Data is saved digitally and can be extracted at any time.

DISADVANTAGES

- Requires that the audience has access to a smartphone or tablet (may be more difficult with younger audiences).

- Needs a stable internet connection.

- There is a cost if access to premium packages is desired.

SURVEYS

DESCRIPTION

Surveys are useful in acquiring specific, structured information and feedback from your audience. Using the same survey both before and after an event allows you to make a direct comparison and gain an idea of the impact on your audience.

Surveys are versatile, they can be quick and provide a snapshot of experiences, or they can be more in-depth and detailed. They therefore vary in terms of how much time is needed for preparation, implementation, and analysis. You can also use a variety of formats and types of questions, e.g., dichotomous, rating scales and open-text responses.

WHAT YOU WILL NEED

- If using an online survey, you will need an account with appropriate software such as Microsoft Forms, Google Forms, Survey Monkey or JISC Online Surveys.
- Allocated time at the start/end of your event to implement the survey.

IMPLEMENTATION

Surveys can be implemented digitally or on paper. Using an online survey platform will save you a lot of time in carrying out your analysis. A paper survey will require you to digitalise the data which can take a long time, particularly when you have a lot of responses. If you are creating your own survey, it is important to test it before you implement it with your audience. Ask a friend or colleague to check it. Check how your survey displays on a phone, tablet, and PC to ensure accessibility. If you can, build the survey into your event, this way, your audience is more likely to complete it than if you send it to them before they arrive at your event or after they have left. Make sure you give your participants sufficient time to complete the survey, this will provide you with more accurate data.

SUMMARY	
Audience Group	Elementary School Pupils High School Pupils College / University Students Teachers Teacher Trainers Families / Communities
Event Type	Classroom Workshop Teacher Training Lectures / Talks / Shows Webinars Website / Blog Competitions / Contents Internships / Placements Social Media Activity Festivals / Celebration Events Astronomy Camps Online Videos
Objective (pupils)	Increase Interest Increase Confidence Increase Future Aspirations
Objective (teachers)	Increase Confidence Increase Astronomy Teaching
When	Before Event During Event After Event
Time	Preparation: Medium Implementation: Low Analysis: Medium

EXAMPLES & TIPS ON DESIGNING QUESTIONS

There are a number of existing instruments that have been designed and tested to measure attitudes and experiences in astronomy and so you may want to adopt one of these existing surveys.

The **Astronomy and Science Student Attitudes (ASSA)** measures eight constructs around students' attitudes towards astronomy and science [6]: interest in astronomy, interest in science outside of school, practical work in science, teacher's actions in science, perception of ability in science, future aspirations in science, the benefits of science, personal relevance of school science. You may want to use the full survey, or just use some of the items to explore specific attitudes. Access the research paper and instrument here.

Two **self-efficacy scales for understanding and doing astronomy** have recently been developed and validated [7]. The two scales measure astronomy personal self-efficacy (identifying personally with understanding astronomy and its concepts.) and instrumental self-efficacy (feelings surrounding adequacy to deal with the instrumentation aspects of using a robotic telescope). You can access the research article and instrument here.

The PhysPort website also has a section dedicated to **'attitude and belief surveys'** which has an array of surveys and provides guidance on how to implement them, the audiences they are suited to, and how long they take to implement. These surveys are less specific to astronomy, focusing on science and physics, however you may find that small adjustments to the questions are sufficient to make them relevant to your programme.

The OAE Centre India has also developed a 16-question baseline survey to explore pupils' understanding of astronomy content in their curriculum, their general interest in astronomy, and connections between their culture knowledge and astronomy. Their survey is available here.

If you would like to design your own survey and questions, Table 9 provides some guidance and tips on how to write a 'good' survey question and some common mistakes to avoid.

TABLE 9 – TIPS AND GUIDANCE FOR DESIGNING SURVEYS

Tip	Explanation
Only ask one thing at a time	An example of a poor question is 'did you enjoy yourself **and** learn something at the event?' You should avoid the word 'and' in your question. What if the participants enjoyed themselves but did not learn anything? Instead, phrase this as two separate questions: 1. Did you enjoy the event? 2. Did you learn anything new at the event?
Avoid leading questions	Leading questions are biased in how they are posed to the participant and *lead* them towards a particular answer. Questions should be neutral in their phrasing. For example, instead of 'how was your experience of our fantastic show today?' try 'how would you describe your experience today?'. You can still ask your participants about particular components of your event, for example, the experiments, using the telescopes, identifying constellations etc., just avoid using adjectives or subjective descriptions in your questions.
Balance positive questions with negative ones	For example, if you ask participants 'what did you like most about the event?', it's good practice to also ask 'what did you like least about the event?'. This prevents you from just collecting positive information and provides a more balanced overview of your participants' experiences.
Consider whether you have any point of comparison or benchmark to compare participants' responses to.	Think about how you will know if you have made a difference through your intervention. Will you ask people questions both before and after they have engaged with your project, or just after? It's important to consider how you phrase questions to make sure you get the information you need. See section 'When to Collect Evaluation Data' for more details.
Include some open questions to give your participants more freedom to express their responses.	A limitation of surveys is that they are rigid and closed questions in particular (where participants must select from pre-defined responses), are susceptible to missing key information. By including open questions that allow participants to provide free-text responses, it provides an opportunity for participants to share information you may not have otherwise captured.
Don't ask more questions than are necessary.	Remember that participants are giving up their time and attention to answer your questions so make sure their time is used wisely. Your aim is to yield as much information as possible from as few questions as you can.
Will your questions elicit data that will map to the objective of your event?	For example, if the objective of your event was to increase your audiences' confidence in astronomy, make sure you're not simply asking questions about their enjoyment of the event, but asking about their sense of understanding, confidence and/or self-efficacy.

ANALYSIS PROCESS

Many online survey sites will do some of the analysis for you. All sites will present quantitative data in relevant bar charts/graphs, and some will even generate word clouds from qualitative data. Depending on what you want to uncover from your data, this level of analysis may be sufficient. For example, if you want to identify what percentage of your audience learnt about a new career pathway at your event, this level of analysis will be adequate. However, if you want to explore differences between groups (e.g. gender), further interrogation will be needed and will involve statistical testing in Microsoft Excel or statistical software (e.g. SPSS, R, Stata). If you collect open-text responses (qualitative data), this can take slightly longer to analyse. One approach is to look for patterns and commonalities in the responses. Have people in your audience responded to your questions in similar ways? What terms or phrases are they using? What overarching themes can you identify from participants' responses? More guidance on how to analyse survey data is provided in Analysing your Data.

ADVANTAGES	DISADVANTAGES
- Easy to implement and can accommodate a large number of responses. - Easy to gain comparable responses. - Relatively straightforward analysis.	- To ensure the quality of a survey, they should be piloted before full implementation. - Surveys can be restrictive – you are limited in the responses you receive by the questions you ask. - To measure change (i.e. the impact of your event) you will need pre- and post-event measurements, or use careful phrasing in your post-survey.

INTERVIEWS

DESCRIPTION

Interviews can be done one-to-one or with a group of participants. They are a great way to capture detailed and contextualised information from your audience. They are much more flexible than surveys and offer more autonomy to the audience in the feedback and experiences they share.

Interviews are particularly useful if you want to explore *why* and *how* questions. Like surveys, interviews can be quick and simple: asking participants two or three questions, or they can be more in-depth: asking a series of in-depth questions and probing participants' responses.

WHAT YOU WILL NEED

- Time allocated to implement data collection and an appropriate space/location.
- Audio recording device.
- A question schedule with some prompts.

IMPLEMENTATION

Interviews are best carried out in a quiet and private space, even if this is just a quiet corner of your venue. Before beginning the interview, you should provide your participants with a summary of the purpose of the exercise and the nature of the questions you will be asking them. Interviews should be audio recorded so you can accurately transcribe participants' comments later. Although it is important to have a set of questions prepared to use as prompts, interviews are a chance to assign more control to the participants and allow them to decide what information they want, or feel is important, to share with you.

SUMMARY	
Audience Group	Elementary School Pupils High School Pupils College / University Students Teachers Teacher Trainers Families / Communities
Event Type	Classroom Workshop Teacher Training Lectures / Talks / Shows Competitions / Contents Internships / Placements Festivals / Celebration Events Astronomy Camps
Objective (pupils)	Increase Understanding Increase Practical Skills Increase Interest Increase Confidence Increase Future Aspirations
Objective (teachers)	Increase Understanding Increase Practical Skills Increase Confidence Improve Teaching Increase Astronomy Teaching
When	Before Event During Event After Event
Time	Preparation: High Implementation: High Analysis: High

EXAMPLES AND TIPS ON DESIGNING QUESTIONS

The types of questions you ask your participants will depend on your objectives and intended outcomes. Some generic questions have been provided below, you may want to use these, or they may give you some inspiration to create some of your own questions. You may also need to adjust the wording of the questions depending on your audience and the context of your programme. However, these may not match your objectives and you may want to write some of your own questions. Table 10 provides some tips on how to write 'good' questions to ask in interviews and some common mistakes to avoid.

General Experiences:
- How would you describe your general experience of today?
- What did you like most about today's event? → What did you like about it?
- What did you like least about today's event? → Why did you not like that?

- How did your experience today compare to your usual experiences of [e.g. science lessons, science centres]?

Future Aspirations
- How do you feel about doing more astronomy in the future? → Why do you feel like that?
- What would you like to be when you grow up? → Why does that career appeal to you?
- How would you feel about having a career in astronomy? → What makes you feel like that?

Feelings and Emotions
- How did you feel during today's event? (prompt: you may have felt happy, sad, excited, frustrated…) → What made you feel like that?

Knowledge and Understanding
- Did you learn anything new today? (if yes), can you explain what you learned?
- What can you tell me about [topic of focus for the session, e.g. Sun-Earth-Moon system]?
- Is there anything you would like to learn more about?

TABLE 10 – TIPS AND GUIDANCE FOR DESIGNING INTERVIEW QUESTION SCHEDULES

Tip	Explanation
Avoid yes/no questions	Aim to ask questions that require a descriptive answer. These kinds of questions start with 'how…', 'in what way…' 'why…' etc. If you need to ask a yes/no question, make sure you have a suitable follow up question that will explore more of the detail surrounding your participant's answer.
Plan your questions ahead	This will help you stay on topic. It also will help you to articulate your questions. If you are not prepared, you may begin to over explain your questions. The more you talk, the less time there is for your participant(s) to talk.
Where appropriate, follow up your participants' answers with 'why?' and 'what' questions.	This will help you understand the reasons behind their experiences and reactions. For example, if they say they would like to pursue a particular career, ask them why they are drawn to that career and what appeals to them. If they say they found something difficult, ask them what in particular they struggled with or what would have helped to make things easier.
Encourage participants to give examples.	The more detail, the better. For example, if the participant said they learnt something new, ask them what they learnt. If they said they enjoyed the presentation, ask them what they liked about it.
Sometimes it pays to be vague.	It may be helpful to ask broad and generic questions, this means that the participants' responses are based on how they interpret the question and what they perceive to be important. However, this does not always work, and some participants will need a bit more structure and prompting than others to help formulate their answers.
Avoid leading questions	Leading questions are biased in how they are posed to the participant and *lead* them towards a particular answer. Questions should be neutral in their phrasing. For example, instead of 'how was your experience of our fantastic show today?' try 'how would you describe your experience today?'. This doesn't mean you can ask your participants about particular components of your event, for example, the experiments, using the telescopes, identifying constellations etc., just avoid using adjectives or subjective descriptions.

Use prompts to encourage your participants or help them formulate their response	As your participant is talking, be encouraging, say things like 'yes' or 'that's interesting', this encourages the participant to continue talking. It's also good to have some more structured prompts in case your participants are struggling to formulate a response. For example, if you asked them 'what career would you like to have when you grow up?', you could follow this with 'would you like to work in an office, or outside, or travelling?' etc.
Do not dominate the conversation.	As a general rule, you as the interviewer should be doing approximately 25% of the talking and your participants should be doing 75% of the talking.
In a group interview, encourage different opinions,	Where someone shares their view or experience, ask others if they agree or had different experiences. This encourages an open conversation. When group members reflect on experiences with one another it can encourage new thinking and understanding.
At the end of the discussion, always ask if they have any other comments they would like to share.	This will provide your participants with some control of what information they share with you and will give you an idea of what they think is important or that they feel strongly about. It also gives you a chance to gain some information you may otherwise have missed. Participants may simply repeat something they have already said but this demonstrates their view of the importance of that point.

ANALYSIS PROCESS

If you are going to the effort of implementing interviews, it is generally recommended you also go to the effort of conducting deeper analysis beyond common words or responses from participants. If you want to do very light touch analysis, it's worth considering using a different data collection approach such as a survey or graffiti wall.

Conversation data can be analysed in a number of ways depending on your objectives and what you want to discover from your data. Thematic analysis is perhaps the most common approach: 'coding' data into different patterns or points of discussion (e.g. likes, dislikes, individuals, STEM topics) and then reviewing and analysing for broader themes and patterns (e.g. effective experiences of engagement, challenges of engagement). More guidance on how to conduct thematic analysis is provided in section Analysing Qualitative Data.

ADVANTAGES	DISADVANTAGES
• Great for capturing detailed information. • They are flexible, you can change your questions as you go along based on participants' responses.	• They are time consuming and hard to do well – practice on friends/colleagues if you can! • They involve small participant numbers. • Can be difficult to recruit participants.

PEER INTERVIEWS

DESCRIPTION

Peer, or friendship interviews are very similar to the method of interviewing but instead of you or another practitioner/evaluator asking the questions, your participants ask the questions to each other. You may find the participants are more relaxed and candid when speaking to their peers, friends, or family, rather than to you as an 'outsider'.

You can do peer interviews in pairs or in groups. A nice approach (particularly with young people) is to have the participants choose some of the questions they want to ask their peers – this can reveal information about what the pupils think is important or interesting to ask about.

WHAT YOU WILL NEED

- Time allocated to implement data collection and an appropriate space/location.
- Audio recording device.
- A question schedule with some prompts.

SUMMARY	
Audience Group	Elementary School Pupils High School Pupils College / University Students Families / Communities
Event Type	Classroom Workshop Competitions / Contents Festivals / Celebration Events Astronomy Camps
Objective (pupils)	Increase Understanding Increase Practical Skills Increase Interest Increase Confidence Increase Future Aspirations
When	Before Event During Event After Event
Time	Preparation: High Implementation: High Analysis: High

IMPLEMENTATION

Peer interviews are best carried out in a quiet and private space, even if this is just a quiet corner of your venue. Before beginning the peer interview, you should provide your participants with a summary of the purpose of the exercise and what you would like them to do. It is important to have some questions prepared ahead of the event and have these on a sheet of paper for the participants to refer to. These questions will help the conversation stay on topic. However, you should encourage your participants to also think of some of their own questions they want to ask their peer(s), this will provide an insight into what information they think is important and interesting. The peer interviews should be audio recorded so you can accurately transcribe participants' comments later. Although it is important to have a set of questions prepared to use as prompts, interviews are a chance to assign more control to the participants and allow them to decide what information they want, or feel is important, to share with you. You can also still be involved in the process too, you may want to have a combination of you and the participants asking questions. If you are involved in the process, you can maintain closer control over the direction of the discourse.

You may want to get creative with how you format the peer interviews, particularly when your participants are young people. You could have them film one another as if they are doing a TV interview, or record a radio set.

EXAMPLES AND SUGGESTIONS

You may want to use the same questions in peer interviews as you would in interviews conducted by you as the evaluator. In which case, take a look at the examples in the 'Interviews' section. However, you could also encourage your participants to think of some of the questions to ask their peers.

ANALYSIS PROCESS

If you are going to the effort of implementing interviews, it is generally recommended you also go to the effort of conducting deeper analysis beyond common words or responses from participants. If you want to do very light touch analysis, it's worth considering using a different data collection approach such as a survey or graffiti wall.

Conversation data can be analysed in a number of ways depending on your objectives and what you want to discover from your data. Thematic analysis is perhaps the most common approach: 'coding' data into different patterns or points of discussion (e.g. likes, dislikes, individuals, STEM topics) and then reviewing and analysing for broader themes and patterns (e.g. effective experiences of engagement, challenges of engagement).

With peer interviews in particular, and where the participants themselves have chosen some of the interview questions, this also presents a valuable source of information. What kind of questions have they chosen to ask their peers? If you have conducted multiple peer interviews, are there any commonalities or differences across what participants choose to ask each other? More guidance on how to conduct thematic analysis is provided in section Analysing Qualitative Data.

ADVANTAGES	DISADVANTAGES
• Great for capturing detailed information. • Assigns more control to the participants and so reveals information about what they think is important to ask about and share. • Participants feel more relaxed when speaking to their peers than with an 'outsider'.	• They are time consuming and hard to do well – practice on friends/colleagues if you can! • They involve small participant numbers. • Can be difficult to recruit participants. • By assigning some control to the participants, discussions can lose focus and digress from your objectives.

JOURNAL / DIARY ENTRIES

DESCRIPTION

Journal / diary entries are best implemented at the end of an event. They involve encouraging your participants to describe and reflect on their experiences of the programme or event, and are particularly well suited to programmes that involve multiple engagements over time. You can provide as much guidance as you want on what participants should contribute in their journal / diary entries.

You can get creative with the format of the journal / diary entries. Entries can be written, recorded as an audio diary, on video or film, or participants could even make scrapbooks and use a combination of images, drawings, and words to summarise their experiences.

WHAT YOU WILL NEED

- Materials for your participants to create their entries (e.g. paper or an appropriate online platform such as Google Forms, Microsoft Forms etc., audio/video devices if recording).
- Allocated time at the end of your session to allow participants to complete their entries.

SUMMARY	
Audience Group	Elementary School Pupils High School Pupils College / University Students Teachers
Event Type	Classroom Workshop Teacher Training Internships / Placements Astronomy Camps
Objective (pupils)	Increase Understanding Increase Interest Increase Confidence Increase Future Aspirations
Objective (teachers)	Increase Understanding Increase Confidence Improve Teaching Increase Astronomy Teaching
When	After Event
Time	Preparation: Low Implementation: Low Analysis: High

IMPLEMENTATION

Journal / diary entries should be implemented at the end of an event, so that participants can reflect on their experiences. Decide ahead of your event what format you want to use to capture your participants' journal /diary entries. You could use a paper form, an online form, audio or video recordings, or a scrapbook. Make sure you provide them with all the material they might need to create their entries.

Journal / diary entries should focus on reflection. Participants start by describing a particular event, reflecting on their experience and involvement during the event, and then consider their feelings, perceptions, and future intentions following the event. Entries can be as open or structured as you like. You should at least provide participants with one or two prompts for what to contribute to their entry. For example, 'describe your experiences of today in as much detail as you can'. Or you might want to give several prompts: What did you do in today's session? How did you feel during today's session? Etc.

For education programmes that involve multiple engagements, participants can complete a journal / diary entry after each session, and you can then explore any changes or progression over the course of the programme.

EXAMPLE PROMPTS

The prompts you provide your participants will depend on the objectives of your programme and your intended outcomes. However, some examples are provided below to get you started. The square brackets are to highlight the type of information these prompts will yield. Figure 5 shows an example diary template.

- Draw an emoji that represents how you felt in today's session. [feelings/emotion]
- What did you do in today's session? [what stood out to them about their experience]
- What do you think you did well at in today's session? [confidence/self-efficacy]
- What did you find most difficult during today's session? [confidence/self-efficacy]
- Is there anything you learnt today that you didn't know yesterday? [learning/understanding]
- What would you like to learn about next? [future intentions]

My Astronomy Diary

Today I felt…

I learnt about…

The best part was…   The worst part was…

--------------------------   --------------------------
--------------------------   --------------------------

Next week, I would like to…   My drawing of today

ANALYSIS PROCESS

How you analyse journal / diary data will depend largely on the format of the data (e.g. writing, drawings, etc.). However, a key focus during the analysis should be on what your audience has decided to reflect on. Reflection involves reviewing one's experience, considering how they feel about it, how it has impacted them and what they plan to do next. What can you learn from their entries about how they are feeling, what they are finding difficult, and what they are learning? What they choose to reflect on in their journal / diary entry will give you an indication of what part of your event was particularly engaging and impactful to your audience (good or bad!). Similarly, it may also be important to consider what is missing from the entries. Was there a particular component of your event that has not been captured in your audiences' reflections?

If you are implementing a programme that involves multiple engagements with your audience over time, you may want to track their entries over the course of the programme. For example, you may want to explore whether they find things more or less difficult as their engagement continues? Do their interests/views/feelings change over time? Do their intentions for the future change over time? Etc.

ADVANTAGES
- Provide insight into individual experiences.
- Can explore understanding, experiences and emotional responses.
- It can be a valuable exercise for your audience to consider their experiences and intentions.
- Can accommodate more accessible formats.

DISADVANTAGES
- Time consuming.
- Can be expensive if you are providing material.
- Less suited to large audience groups.

TESTS / CONCEPT INVENTORIES

DESCRIPTION

Where surveys capture attitudes and experiences, tests and quizzes capture knowledge and understanding. You can use the same online software as you would for surveys. They are therefore a great way to gauge whether you have been successful in promoting awareness, knowledge or understanding, and perhaps where there remains room for improvement.

Tests are most effective when you know your audiences' 'baseline' understanding and therefore it can be helpful to implement parallel tests both before and after your event.

WHAT YOU WILL NEED

- If using an online survey, you will need an account with appropriate software such as Microsoft Forms, Google Forms, Survey Monkey or Online Surveys.
- Allocated time at the start/end of your event to implement the test.

IMPLEMENTATION

Tests / concept inventories can be implemented digitally or on paper. Using an online survey platform will save you a lot of time in carrying out your analysis. A paper test will require you to digitalise the data which can take a long time, particularly when you have a lot of responses. If you are creating your own test, it is important to trial it before you implement it with your audience. Ask a friend

SUMMARY	
Audience Group	Elementary School Pupils High School Pupils College / University Students Teachers Teacher Trainers Families / Communities
Event Type	Classroom Workshop Teacher Training Lectures / Talks / Shows Webinars Website / Blog Competitions / Contents Internships / Placements Social Media Activity Festivals / Celebration Events Astronomy Camps Online Videos
Objective (pupils)	Increase Understanding Increase Practical Skills
Objective (teachers)	Increase Understanding Increase Practical Skills
When	Before Event During Event After Event
Time	Preparation: Medium Implementation: Low Analysis: Medium

or colleague to check it. Check how your test displays on a phone, tablet, and PC to ensure accessibility. If you can, build the test into your event, this way, your audience is more likely to complete it than if you send it to them before they arrive at your event or after they have left. Make sure you give your participants sufficient time to complete the test, this will provide you with more accurate data.

EXAMPLE TESTS & INVENTORIES

The PhysPort website provides an array of existing concept inventories that have been implemented with different audiences. You can search for concept inventories/tests based on topic, education level, format, and language translations. It is difficult to make suggestions for questions you might want to include in your tests and they will depend on the content of your programme. However, it is important that the questions are relevant to your programme and suitable for your audience.

ANALYSIS PROCESS

Many online test/survey sites will do much of the analysis for you as you can specify the 'correct' answers. This will of course only work for closed, multiple choice questions, not for questions that ask for open text responses. All sites will give you an overview of the results and the number/proportion of correct answers. Depending on what you want to uncover from your data, this level of analysis may be sufficient. For example, if you want to

identify what percentage of your audience were able to list the planets in order of closest to furthest distance from the Sun, this level of analysis will be adequate. Using existing concept inventories, such as those on the PhysPort site will often also provide you a benchmark of other data to compare to. However, if you want to explore differences between groups within your own sample (e.g. by gender), further interrogation will be needed and will involve statistical testing in Microsoft Excel or statistical software (e.g. SPSS, R, Stata). This software will also allow you to conduct timepoint comparisons, i.e., comparing before and after results to explore whether your participants' test scores increased after your event. You can also look at specific questions within your test/quiz: if participants make a big improvement in their understanding in one area but no improvement in another, you know which part of your event might need revising or a greater focus in future. More guidance on how to analyse test data is provided in Analysing Quantitative Data.

ADVANTAGES	DISADVANTAGES
• Relatively straightforward analysis	• Can take the 'fun' out of an event – sense of being tested is not always welcomed by audiences
• Easy to implement and can accommodate a large number of responses	• To measure gained understanding (i.e. impact of your event) you will need pre- and post-event measurements

DRAWING ACTIVITY

DESCRIPTION

Drawing is a fun, accessible activity that can reveal a lot of information about your audience. The most famous example of this in practice is the 'draw a scientist' activity that revealed a wealth of information about stereotypes and young peoples' perception of scientists. When asked to draw something related to themselves, people's drawings can reveal a lot about their identities, emotions, and experiences.

WHAT YOU WILL NEED

- Designated space with paper and pens.
- Some prompts for what you would like your audience to draw.
- If implementing online, you'll need a relevant software account, e.g., Mural, Miro.

IMPLEMENTATION

It is important you guide your audience in what kind of

SUMMARY	
Audience Group	Elementary School Pupils (young) High School Pupils Families / Communities
Event Type	Classroom Workshop Lectures / Talks / Shows Festivals / Celebration Events Astronomy Camps
Objective (pupils)	Increase Understanding Increase Interest Increase Future Aspirations
When	Before Event During Event After Event
Time	Preparation: Low Implementation: Medium Analysis: High

things you want them to draw. This should also correspond with the objective of your programme. If your objective is to promote understanding of careers in astronomy, then you could ask them to draw an astronomer, but you wouldn't ask them to draw a space rocket. Depending on your objectives, you could ask your audience to draw what they would like to be when they grow up (aspirations), draw themselves doing astronomy (identity in astronomy), draw your favourite part of today's event (experience), draw an astronomer (understanding). You can also use a drawing exercise to explore your audience's understanding of astronomical concepts, for example, you could ask them to draw a diagram of the Earth-Moon-Sun system. When working with young people, make sure you set aside sufficient time in your event for them to complete their drawings. Some may want to take them home with them, if this is the case, ask if you can take a photograph of their drawing.

EXAMPLE DRAWING PROMPTS

Below are some example prompts and questions you may want to use to encourage your audience's drawings. You may want to adopt these, or they may simply give you some inspiration for more suitable prompts for your own programme and your specific objectives.

Attitudes and Identities:

- Draw an emoji that represents how you felt during today's session.
- Draw a picture of you during today's session.
- Draw a picture of your favourite thing about today's session.
- Draw a picture of today's session.
- Draw an astronomer at work.
- Draw what you would like to be when you grow up.

Knowledge and Understanding:

- Draw a diagram of our Solar System.
- Draw the different phases of the Moon.
- Draw the lifecycle of a solar mass star.

ANALYSIS PROCESS

Analysis of drawings can be approached in a similar way to analysis of qualitative data (interviews, open-ended questions, etc). A good way to start is to look through all the drawings to get a sense of any patterns or themes among them. When you get a sense of the general content of the drawings, you can start categorising them into these patterns. For example, if you asked your audience to draw what they would like to be when they grow up, are there any common careers such as desk-based jobs, active jobs, do they relate to particular subjects? Are there any differences between drawings from males and females? Are any of the careers relevant to your event? If you asked them to draw themselves doing science, what are they doing? Are they by themselves or with their peers? Are they doing an experiment or working at a desk or computer? What facial expression have they given themselves? There are a variety of things you can look for in drawing data and you may even want to explore more than one of these examples provided. More guidance on how to analyse your data is provided in Analysing Qualitative Data.

ADVANTAGES	DISADVANTAGES
• It's a fun activity for your audience and particularly when working with children.	• It's not for everyone – this really works best with young children (e.g. under the age of 15)
• Drawing is a more accessible and less intimidating activity than interviewing, surveys, or tests.	• Analysis can be tricky, it's important to provide your audience with a prompt of what they should draw and consider what you want to look for in your analysis when choosing your prompt.
	• Time consuming.

Analysing your Data

Analysis can be a very lengthy, in-depth process that often requires careful consideration of epistemologies, theoretical lenses, validity, and reliability. It is important to remember here that we are talking about evaluation, not research. Although your evaluations may lead to publication in research and academic journals, the purpose of this document is not to guide you through conducting astronomy education research. The purpose is to help you formulate appropriate evaluation plans and models for your astronomy education programme, identify appropriate evaluation data collection tools, and to then make sense of the data once you have it.

This section therefore provides an overview of some of the common analysis techniques that are used with the type of data you will be collecting. This is intended to give you an idea of what you can do with your data once you have it.

Analysing Qualitative Data

Regardless of objectives and analysis approach, the collection of qualitative data will often involve some element of coding. Coding describes the process of identifying a feature of your data that appears interesting [8], for example, that relates to your programme objectives and intended outcomes.

Coding involves assigning a label to a piece of text to describe the key idea or information you feel that text represents [9]. The 'label' is the code itself and should clearly encapsulate what that piece of information describes. For example, the statement 'I enjoyed the observing night and seeing Jupiter through the telescope' might be coded as 'enjoyment'.

Coding helps you to organise and structure your data, and identify patterns of similar information across the data. Coding is therefore sometimes considered akin to indexing or categorising of information [9]. You should not worry about finding the perfect code for each piece of data, you can assign data to more than one code, change your codes, combine codes, and so on. Coding is an ongoing, iterative process. For example, the statement above might be coded as 'enjoyment' but may also be coded as 'action - observing'. However, although your data might apply to multiple codes, each code should be discrete from another code, you should avoid having multiple codes that describe the same feature in the data. For example, it is unlikely you would have a code for 'enjoyment' and a code for 'fun'.

Figure 6 provides an example excerpt from an interview with a school pupil who attended an observing night. The text has been colour coordinated to demonstrate how you might code this text.

FIGURE 6 – EXAMPLE CODING OF QUALITATIVE DATA FROM AN INTERVIEW EXCERPT
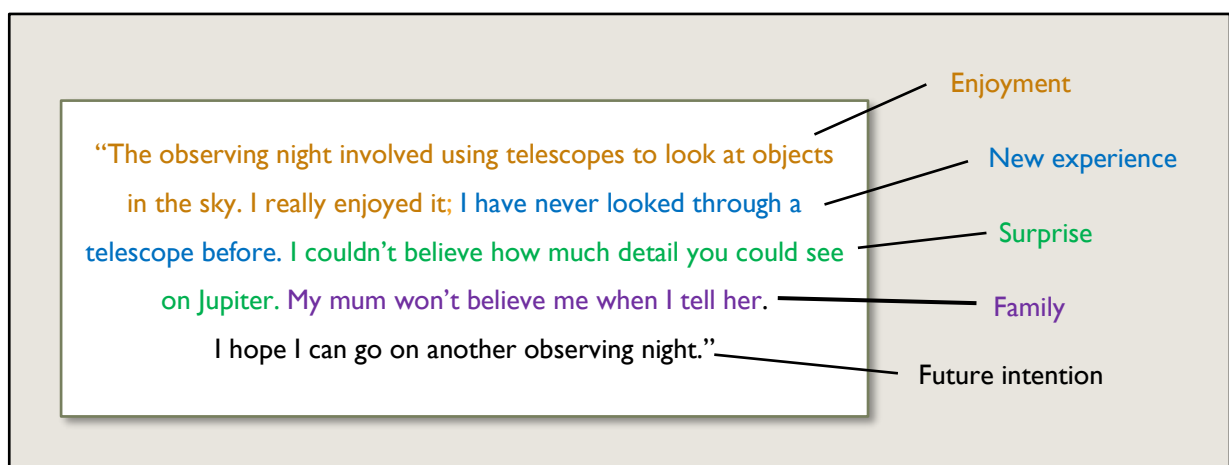
Once you have your codes, you can start to look for patterns, connections, and relationships across your codes. A widely used and accessible process is thematic analysis. Thematic analysis is a method used to identify, analyse, and report patterns and themes within a dataset [8]. A theme is different to a code in that a theme should be applicable across multiple codes and relate directly to your objectives. Themes are a way to explain relationships between your codes and how they map to your objectives. These themes are often used to structure how you report on your analysis and your findings.

Braun and Clarke [8] set out several key steps for conducting thematic analysis and are summarised in Table 11.
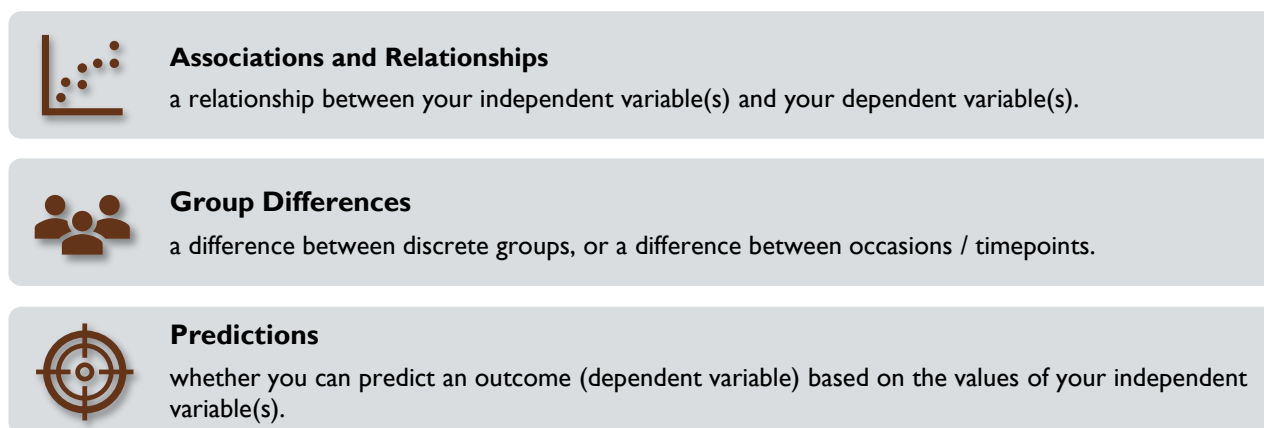
TABLE 11 – STEPS FOR CONDUCTING THEMATIC ANALYSIS ADAPTED FROM BRAUN AND CLARKE [8]

Step	Description
Familiarising yourself with your data	Make sure you have your data organised into a manageable format, e.g., a transcription document, a file of drawings, a document of text etc. Look/read through this data to get an understanding of its overall content.
Allocating codes	Assign labels to your data that describe the key idea or information you feel that data represents. It may not be appropriate to code every piece of data, often you will yield surplus data that is not relevant to your objective. However, you should review your entire dataset during the coding process.
Constructing your themes	Review your codes to see what patterns and themes can describe the codes you have allocated to your data. As you begin to construct the themes of your data you should identify all the codes that are relevant to these themes.
Reviewing and defining your themes	It may be helpful to create a 'map' of your themes, their relevant codes, and your objectives to check they are coherent, appropriate, and relevant to your project. Here you can also refine your themes and give them a name and definition to help describe the 'story' of your data.
Compiling your report	Set out the 'story' or 'narrative' of your data and of your findings. Here you can select particular extracts from your dataset (e.g. a quote, an observation, a participants' drawing) to demonstrate your findings.

Analysing Quantitative Data

There are many different statistical tests that can be used to analyse quantitative data. To identify the most appropriate statistical test for the data you have, the first thing to consider is what you are looking for in your data and what you intend to report on. This ties back to your objective and your intended outcomes. Generally speaking, there are three key elements you can explore with quantitative data, these are summarised in Figure 7.

FIGURE 7 – WHAT YOU CAN EXPLORE THROUGH QUANTITATIVE DATA

Associations and Relationships
a relationship between your independent variable(s) and your dependent variable(s).

Group Differences
a difference between discrete groups, or a difference between occasions / timepoints.

Predictions
whether you can predict an outcome (dependent variable) based on the values of your independent variable(s).

Unless you have got quantitative data from ~100 participants or more, and you are looking to publish your data in academic journals, it is unlikely that you will need to look at tests that allow you to make predictions. Predictions are essentially a step beyond the tests that look for associations and relationships but have been included here as they are important to recognise.

It is important to highlight that this toolkit is not intended to provide a comprehensive course in statistical analysis, but rather provide you with a starting point to know what you can do with your quantitative data and how to make sense of it.

The next key factor in determining which statistical test you should use is what type of data you have. Quantitative data can have several different formats, categorical, ordinal, or continuous. Their key characteristics are summarised in Table 12.

TABLE 12 – TYPES OF QUANTITATIVE DATA

Type of data	Description
Categorical	Categorical data has no particular order, they are discrete from one another. Examples include subjects (e.g. mathematics, physics, geography), countries, gender, etc. Categorical data may also be termed discrete or nominal data, or dichotomous if there are only two options (e.g. pass/fail).
Ordinal	Ordinal data are similar to categorical data but can be ordered or ranked. The most common examples are Likert scales, where participants are asked to respond to a particular statement by selecting one of several options. For example, participants respond to the statement 'I enjoy astronomy' with either 'strongly disagree', 'disagree', 'neither', 'agree', or 'strongly agree'. Or, for the question 'how helpful/unhelpful did you find today's lesson'? participants can select 'very unhelpful', 'quite unhelpful' 'neither', 'quite helpful', or 'very helpful'.
Continuous	Continuous variables provide data that can be measured on a continuum. Their values are truly numerical, for example age, test scores, number of engagements.

Once you know what you want to look for in your data, and the type of data you are working with, you can explore some of the relevant statistical tests. Statistical tests still require certain assumptions to be met, this includes factors such as the distribution of your data, how many variables you have, whether you have discrete or related groups, and so on. Table 13 summarises some of these tests and the parameters your data should meet for the test to be appropriate.

TABLE 13 – STATISTICAL TESTS AND THEIR REQUIREMENTS

What are you looking for in your data?	The types of statistical tests you can do	When the test is appropriate
Associations and relationships	Chi-square test of association	You have **categorical** data.
	Pearson's R correlation	You have **continuous** data that shows a **normal distribution**.
	Spearman's Rank correlation	You have **continuous** data that does **not show a normal distribution.**
Group differences	Chi-square test of independence	You have **categorical** data.
	ANOVA (analysis of variance)	You have **continuous** data, are **comparing three groups or more** (either discrete groups, or the same group over multiple occasions) and your data shows a **normal distribution.**

	Kruskal Wallis test	You have **ordinal** data (or continuous data that does **not show a normal distribution**), are **comparing three or more discrete groups**.
	Friedman test	You have **ordinal** data (or continuous data that does **not show a normal distribution**), and are comparing the same group but over **three or more occasions**.
	Independent samples t-test	You have **continuous** data, are comparing **two discrete groups**, and your data shows a **normal distribution**.
	Mann Whitney U test	You have **ordinal** data (or continuous data that does **not show a normal distribution**) and are comparing **two discrete** groups.
	Paired samples t-test	You have **continuous** data, are comparing the same group but across **two occasions**, and your data shows a **normal distribution**.
	Wilcoxon Signed Rank test	You have **ordinal** data (or continuous data that does **not show a normal distribution**) and are comparing the same group across **two occasions**.
Predictions	Linear regression	You want to predict the value of a **continuous** dependent variable, based on the value of a **single continuous** independent variable.
	Ordinal logistic regression	You want to predict the value of an **ordinal** dependent variable based on your independent variable(s).
	Binomial regression	You want to predict a **dichotomous** dependent variable based on your independent variable(s).
	Multiple regression	You want to predict the value of a **continuous** dependent variable, based on the value of **multiple continuous** independent variables.

EVALUATION DATA REPOSITORY

The OAE Evaluation Data Repository is where you will submit the evaluation data you have collected from your astronomy education events. Before you submit your data, you will be asked to complete an online form that provides a series of questions intended to capture the contextual elements of your astronomy education event. All NAECs will be asked to provide this so we can begin to build a global picture of astronomy education initiatives across our network and understand what practices work under what circumstances.

The evaluation data repository can be accessed here.

REFERENCES

1. Pawson R. and Tilley N. 1997. *Realistic Evaluation*. London: Sage

2. Greenhalgh J. and Manzano A. 2022. Understanding 'context' in realist evaluation and synthesis. *International Journal of Social Research Methodology* 25(5), pp. 583-595. DOI: 10.1080/ 3645579.2021.1918484

3. Rameses II Project. 2017. *What realists mean by context*. [Online] Available at: https://www.ramesesproject.org/

4. Rameses II Project. 2017. *What is a mechanism?* [Online] Available at: https://www.ramesesproject.org/

5. Better Evaluation. 2021. *Realist Evaluation*. [Online] Available at: https://www.betterevaluation.org/methods-approaches/approaches/realist-evaluation

6. Bartlett S, Fitzgerald MT, McKinnon DH, Danaia L, and Lazendic-Galloway J. 2018. Astronomy and Science Student Attitudes (ASSA): a short review and validation of a new instrument. *Journal of Astronomy & Earth Sciences Education (JAESE)* 5(1). DOI: 10.19030/jaese.v5i1.10190

7. Freed R, McKinnon D, Fitzgerald M, and Norris CM. 2022. Development and validation of an astronomy self-efficacy instrument for understanding and doing. *Physical Review Physics Education Research* 18. DOI: 10.1103/PhysRevPhysEducRes.18.010117

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

9. Cohen L, Manion L, and Morrison K. 2018. *Research Methods in Education*. 8th Edition. Oxfordshire: Routledge

https://astro4edu.org/

To cite this document: