2025

EVALUATION REVIEW

JANUARY 2025

SOPHIE BARTLETT

Contents

BACKGROUND

There are myriad astronomy education initiatives, activities, programmes and projects taking place across the field and at various levels. From bespoke lesson plans, teacher training courses, astronomy Olympiads, and intensive astronomy camps. We also see a variety of overarching objectives, whether they be to demonstrate the value of astronomy to promote its inclusion in the curriculum, enhancing the teaching of astronomy as its own subject, or using astronomy as a point of engagement for other subjects.

To ensure that all of these endeavours stand the best chance of success, requires thoughtful reflection, evidence-based decision making and ongoing improvements. Evaluation is paramount to achieving this and provides a systematic approach to understanding what works, what doesn't work, where and why.

This document has been created to provide a foundation for astronomy educators to 'demystify' evaluation and guide them through the evaluation process, providing clear and practical insights into principles, methods and applications of evaluation in a variety of settings.

Purpose of this Document

The aim of this document is five-fold:

1. **Inspire Astronomy Educators**: equip astronomy educators with the background knowledge and examples to implement meaningful evaluations of their astronomy education activities, projects and programmes.

2. **Promote a Culture that Embraces Evaluation Practice**: encourage the astronomy education field to embrace a culture where evaluation is commonplace and embedded in all activities, rather than a 'nice to have' or an afterthought.

3. **Build a Shared Understanding**: work towards establishing a common language and culture of evaluation to ensure consistency, clarity and value in its use.

4. **Encourage Ongoing Improvement**: foster a culture within astronomy education that strives for continuous development by harnessing evaluation to identify strengths, challenges and gaps in astronomy education practices.

5. **Bridge Theory and Practice**: provide practical examples and concrete pathways to apply evaluation theories to real-world applications.

AN INTRODUCTION TO EVALUATION

The term 'evaluation' can be attributed to multiple definitions depending on setting, context, cultures and countries. The OAE considers evaluation to be a process that captures information around the value, efficacy, effectiveness and impact of a programme, project or initiative. It uncovers how well something works and the extent to which the intended outcomes were achieved. In educational settings, evaluation is often used to determine how well pupils, teachers, trainers, programmes or schools are achieving their intended goals and improving learning or attitudinal outcomes.

Evaluation allows us to answer questions such as: to what extent did our initiative achieve its goals? How can our programme be improved? Should a project continue? What is the impact of our initiatives on our audiences? Whether these be questions about teacher training events, classroom interventions, a new curriculum or a single resource, evaluation offers valuable insights.

Types of Evaluation

Evaluation can serve multiple purposes and is able to suit a variety of needs. As educators, we are always looking to improve, sustain or upscale education initiatives. Whatever your main objective is, an appropriate evaluation protocol will help you identify the extent to which you are achieving your objectives.

There are generally three broad types of evaluation that vary depending on when they are conducted in relation to a project: diagnostic, formative, and summative evaluation. Each of these are briefly summarised below.

Diagnostic Evaluation	Diagnostic evaluation is helpful when you have **limited information to begin with**. It takes place at the start of a project to **identify the needs of your audience** / participants, their strengths and weaknesses, resources and obstacles, so you can **effectively tailor your initiative** to suit their needs.
Formative Evaluation	Formative evaluation captures information about a project or initiative while it is still **in development**. This helps you to identify components that are **working well**, and components that **need improving or changing**. Formative evaluation is therefore an ongoing or repeated process that **informs the development of the final version** of your project.
Summative Evaluation	Summative evaluation allows you to identify whether you **achieved what you set out to do**. It determines the outcomes and **overall effectiveness** of a project. This can help in justifying the continuation or upscaling of the initiative, or where negative results are drawn, what should be changed in future.

Why is Evaluation Helpful?

Evaluation is helpful in acquiring an evidence-base to qualify your practice, decisions or claims. There are myriad benefits of conducting evaluation, but a few examples are provided below:

Inform Decision-Making: results from an evaluation can help you make important decisions about your activities. For example, diagnostic evaluation could help you identify a target audience that would most benefit from your project. Or alternatively, if you know who your target audience is, diagnostic evaluation can help you design a project most suited to their needs.

Understand your Audience: like the examples for informing decision-making, evaluating a project can help you to learn more about your target audience. For example, you may learn more about their interests, their preferred learning styles, or their career aspirations. This information can help you improve your engagement with this audience.

Enhance your Accountability: documentation of the implementation process of your project and your evaluation protocols provides an evidence-base for how you allocated your resources, how you made decisions, and how implementation took place. This information can be helpful to share with funders or relevant stakeholders such as governing bodies, parents, headteachers etc.

Give Confidence to your Outcomes: we often rely on anecdotes or comments from audience members, but data collected through a systematic evaluation provides greater assurance, to you and your stakeholders, in whether or not your project was successful, whether your objectives were achieved, and what to do next.

Justify Funding: funding for educational initiatives is competitive so it is important you are able to demonstrate your integrity and competency in delivering on your proposal. Evidence of careful consideration of how you will evaluate your project will be welcomed by reviewers, or evidence from evaluation of a previous a pilot or similar initiative that you can report on, will give them confidence in your capability to deliver the project.

Improve Efficiency: sometimes a project can have a positive outcome, but what led to such an outcome might be unclear. For example, if a two-day workshop with teachers improved their understanding, it may be that it was just a two-hour activity that led to this improvement, and you could have done the workshop in half a day, rather than two days. Evaluating the full workshop will help you identify what elements of your project are working well, and which are not, enabling you to streamline your programmes and implement them most efficiently.

Publicise your Findings: if you have systematically evaluated a project, then it is worth considering whether an academic journal or educators' magazine might be interested in publishing your results. Such a publication is helpful for disseminating your work and developing your own portfolio.

REVIEW

Before conducting an evaluation, it is helpful to learn about what evaluation practices have been implemented elsewhere with similar objectives. A literature review was therefore carried out to explore astronomy education evaluation practices carried out in the past ten years, and the methodologies employed. The review was intended as a scoping exercise with a purpose to identify what had been done. It was therefore purposefully inclusive and did not involve quality assessment criteria as in a traditional literature review. Instead, the objectives of this review were four-fold:

1. To explore what evaluation studies have been conducted in the field of astronomy education at primary and secondary levels of education (pre-tertiary).
2. To identify common focus areas, research questions or evaluation objectives among the studies.
3. To identify common study designs and methods for conducting astronomy education evaluation studies.
4. To provide an overview of evaluation methods to astronomy educators new to evaluation.

The studies extracted for review are by no means an exhaustive list of astronomy education activities and programmes that have been evaluated in the past ten years. Many will have been conducted in-house and not made publicly available, others will be in the form of final reports or deliverables to funders not listed in databases, and some may be available through project or institutional websites. A full comprehensive search was beyond the scope of this document and would have required substantial resource to manage the volume of literature. Instead, the evaluation studies reported here, draw on those that could by systematically searched and retrieved through academic research databases.

Five databases were searched: Scopus[1], Web of Science[2], ProQuest[3], EBSCOHost[4] (including: BEI, Communication Abstracts, ERIC, Education Abstracts, Educational Administration Abstracts, Child Development and Adolescent Studies), and Istardb[5]. An advanced search with Boolean operators was used for each database apart from Istardb where this was not permissible. Search parameters were as follows:

Title (astronomy OR cosmology OR astrophysics)

AND Title (education OR teaching OR initiative OR program* OR activity OR lesson OR practices OR experience OR project OR intervention OR training)

AND Title-Abstract (evaluat*)

These parameters alone retrieved several hundred articles, reports and theses; therefore, the search was further refined to only include material published in the last 10 years (2014-2024), in English language, and full texts (not just abstracts). Full texts were not limited to journal articles but also included published conference proceedings, dissertations/theses, and published reports. This yielded 154 studies, 60 of which were duplicates. After duplicates were removed, titles and abstracts were screened for the following criteria: relevance, primary research, compulsory stages of education/teaching, this left 52 studies for review. A PRISMA flow diagram of the process is provided in Figure 1.

[1] https://www.scopus.com/home.uri
[2] https://www.webofscience.com/wos/woscc/basic-search
[3] https://www.proquest.com/
[4] https://login.ebsco.com/
[5] https://istardb.org/P

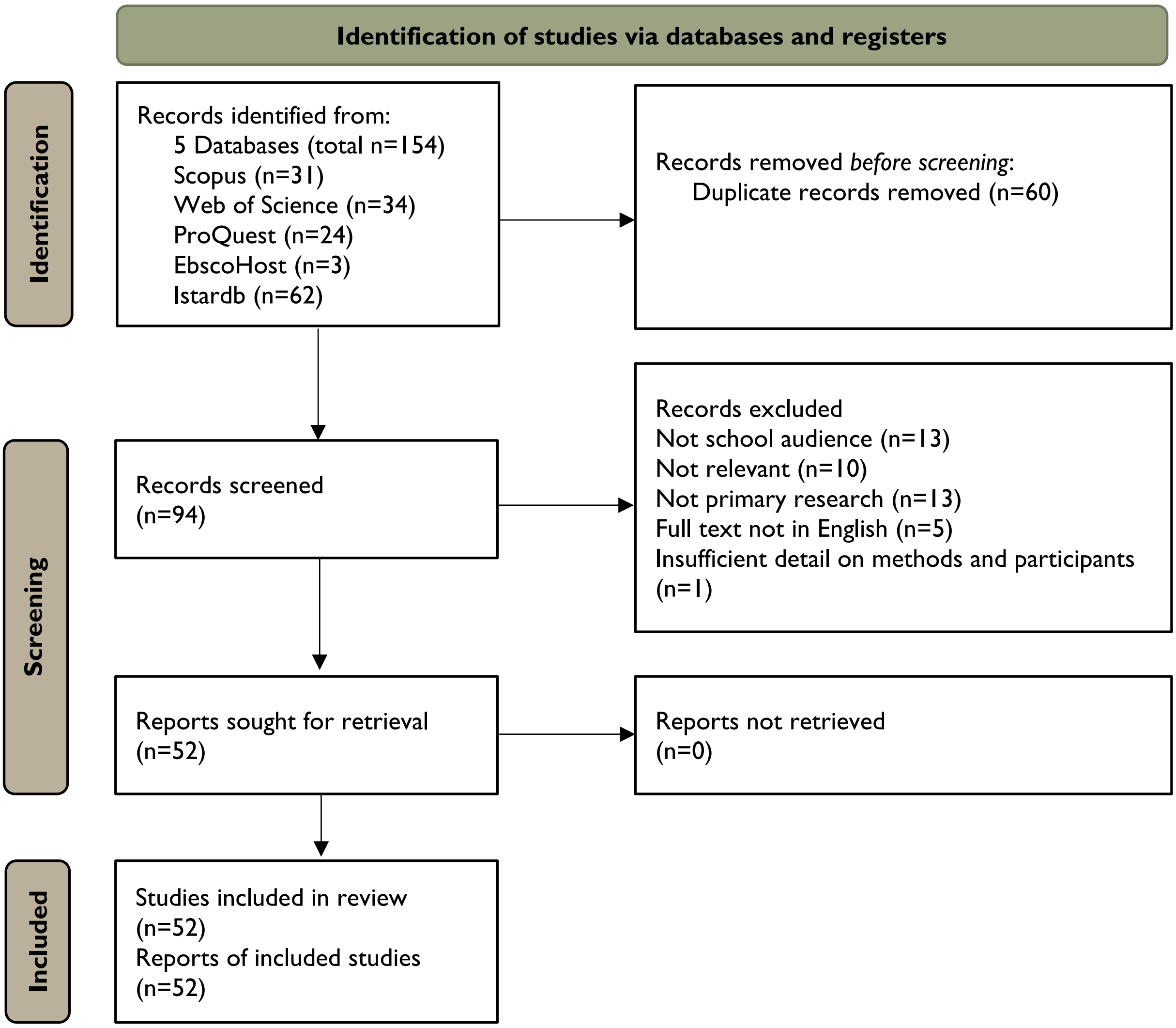

FIGURE 1: PRISMA FLOW DIAGRAM OF LITERATURE SEARCH

In total, 52 evaluation studies were reviewed. As this endeavour was intended to be a status review; an identification of what has been done, a quality assessment was not necessary. Instead, each of the 52 studies were reviewed to identify the key components surrounding the focus, methods and analyses of the studies, this included:

- Focus of evaluation
- Educational setting: classroom, teacher training, informal, other
- Participants: primary school pupils, secondary school pupils, primary school teachers, secondary school teachers, teacher trainers
- Evaluation method: quantitative, qualitative, mixed methods
- Evaluation design
- Evaluation questions/objectives
- Data collection tools
- Use of previously developed data collection instruments
- Mention of reliability or validity measures
- Data analysis techniques
- Key findings
- Limitations

The following sections describe the key features of the 52 studies to provide the reader with examples of different types of studies and the types of findings they were able to yield to signpost the reader to relevant published studies that may be useful in informing their own astronomy education evaluation projects. These key features are structured around three sub-headings: focus of evaluation, methods, and study designs.

Focus of Evaluation

The studies largely focused on one or more four key areas: knowledge, attitude, experience, or data collection tool development. Most prevalent among the studies was astronomy knowledge, which was a focus in 37 studies, however 23 studies had more than one focus, for example explored both experience and knowledge, or knowledge and attitude.

Knowledge-focus

The knowledge-focused studies mostly centred around improving pupils' or teachers' astronomy knowledge and understanding, or improving teachers' pedagogical content knowledge for teaching astronomy. In terms of pupils' astronomy knowledge, some studies evaluated the effectiveness of a particular astronomy education programme (Dankenbring & Capobianco, 2016; Türk et al., 2015). For example, (Türk et al., 2015) explored the impact of a five-day astronomy summer project involving hands-on activities, outdoor practices, a planetarium visit and observatory activities on both pupils' and teachers' astronomy knowledge. Comparison of astronomy knowledge test scores before and after the summer programme revealed significant increases in test scores for both pupils and teachers. Other studies explored particular teaching methods in the classroom and their impact on performance or understanding of astronomical concepts. One study investigated different technology-enhanced learning environments on pupils' astronomy knowledge retention (Zimmerman et al., 2014). Elsewhere, real and virtual models were evaluated and compared in terms of their effectiveness in promoting secondary school pupils' understanding of Earth-Sun-Moon motions (Tsihouridis et al., 2024). Pupils' understanding of solar eclipses saw the biggest improvements from both models, however the virtual models helped pupils to observe and clarify features through better visualisation.

The value of game-based learning in astronomy has been explored both in terms of board games (Cardinot & Fairfield, 2019a), and digital games (Liu et al., 2014). In both cases, pupils' astronomy knowledge increased significantly, this was seen for primary school pupils in the evaluation of a digital game and secondary school pupils for the evaluation of a board game. However, neither evaluation study involved control groups or compared these game-based environments to other learning environments. Learning through storytelling has also been seen to improve astronomy literacy among secondary school pupils for whom English is a second/foreign language (Chubko et al., 2019).

Several studies have explored the role of secondary school pupils' learning through authentic science experiences. One study found that engaging in the scientific process; designing investigations, conducting analyses and constructing argumentation, enhanced pupils' scientific skills significantly more than a control group (Kuhn et al., 2017). Elsewhere, engagement with the scientific database, Aladin Sky Atlas[6], to search for, view, and manipulate astronomical images significantly enhanced pupils' understanding of astronomical concepts and objects (de Lima et al., 2018). Pupils' engagement in science practices has also been explored among younger pupils. Plummer, Schultz, et al. (2015) found that pre-school children were capable of

[6] https://aladin.cds.unistra.fr/

engaging in early stages of scientific practices and teachers can support this through providing opportunities for first-hand experiences of science through physical models, observations and evidence collection.

Authentic science experiences or scientific-inquiry have also been evaluated in terms of their impact on teachers. A mixed method study with primary and secondary school teachers revealed that while teachers' knowledge scores improved between pre- and post-tests, some struggled with the open-ended nature of scientific inquiry and desired more structured teaching content and outputs (Burrows et al., 2016). However, another study involving teacher engagement in astronomy research revealed that ongoing participation fostered teachers' links to the research community which prompted increasing comfort and confidence in the unknown (Rebull, French, et al., 2018). Lewis (2019) coined the term 'practice what you teach', demonstrating the value of pre-service teacher training that models science lessons; teaching through inquiry, rather than teaching about inquiry.

At a more micro level, some evaluation studies investigated the implementation of specific classroom resources or modules. For example, implementation of an augmented-reality tool was most effective in knowledge retention among primary school pupils when coupled with outdoor teaching. This retention was seen both immediately and 30 days subsequent to engagement (Zhang et al., 2014). Similarly, a 14-item astronomy concepts pre- and post-test was implemented with two groups of secondary school pupils, one control group with traditional classroom teaching, and one experimental group who experienced a planetarium show. Those who experienced the planetarium show improved significantly more on eight of the 14 questions than the control group (Türk & Kalkan, 2015).

Some studies did not look specifically at outcomes of an intervention or implementation but explored the current status of understanding among participants. For example, understanding secondary school pupils' level of understanding of the formation of the Solar System (Plummer et al., 2015), or understanding of celestial motion in social and cultural contexts (Anantasook et al., 2015). Some also looked at teachers' understanding, for example their beliefs in pseudo-science (Kaplan, 2014) or understanding of astronomy content in national curricula (Rodrigues de Andrade, 2023). Such studies could be considered as diagnostic evaluation that set out to assess the current state of a particular audience's understanding of astronomy that could inform recommendations or actions for teaching or intervention programmes.

In the case of teachers, some studies evaluated the impact of a teacher training event on teachers' astronomy knowledge. Evaluation of a blended (online and in-person) training programme for primary school teachers demonstrated the value of peer interactions among teachers to consolidate practices, it also revealed the challenges teachers face in delivering age-appropriate classroom activities and refraining from teacher-directed practices (Ampartzaki et al., 2024).

Experience-focus

Twenty studies evaluated pupils' and/or teachers' experiences of participating and engaging in a particular astronomy education programme or activity. These were not necessarily always about the resulting impact of these programmes, but also the particular experiences of participants during their engagement. In Australia, several studies have evaluated the role of indigenous and aboriginal content in astronomy learning. In an aim to make Aboriginal astronomy relevant beyond just Aboriginal pupils, one study (Wyatt et al., 2014), revealed the importance of using specifically local Aboriginal content rather than wider country-relevant content. Nonetheless, the importance of not detracting from other valuable learning experiences such as engagement in hands-on activities. Elsewhere, a primary school case study demonstrated how indigenous storytelling through interaction with local Aboriginal elders and community members enabled pupils, approximately half of whom were Indigenous, to make personal connections with the storytellers as well as identify connections and differences with such stories and science in Western cultures (Ruddell et al., 2016).

There have been several studies that evaluated the implementation of bespoke resources for pupils with disabilities or additional learning needs. For example, subsequent to engagement with sonification tools, partially sighted pupils reported positive perceptions on the accessibility of astronomy and their ability to engage with astronomy content (Guiotto et al., 2023). Alternatively, ongoing formative evaluation of "the vibrating universe" project revealed valuable information around the challenges in creating suitable resources for deaf pupils (De Leo-Winkler et al., 2019).

Several studies explored pupils' and teachers' experiences of engaging with digital technologies. One explored both pupils and teachers' perspectives of using 3D technologies in the classroom (Isik-Ercan et al., 2014). While this evaluation had both a knowledge and experience focus, the latter focus revealed that pupils were excited by the novelty of these technologies which created an increased motivation to learn. Teachers also regarded 3D technologies as a useful complimentary tool for other teaching approaches rather than as a sole primary approach. Elsewhere, an AR mobile app as a tactile learning tool was praised by teachers in terms of its capacity for personalisation and its offering of autonomy to pupils in creating their own learning pathway (Antoniou et al., 2018).

One study investigated how teachers select particular professional development programmes and their prior experiences in order to identify the needs of these audiences (Rebull, Roberts, et al., 2018). In their programme involving training teachers in conducting authentic astronomy research, Rebull et al. identified one group that had the relevant skills and background knowledge but were new to the process itself, others were overqualified so had little to learn from the programme, some did not have any working knowledge of astronomical concepts so needed to be brought up to speed, and others viewed the programme as a way to fulfil their professional development commitment rather than due to innate enthusiasm for the topic.

Attitude-focus

Of the literature reviewed, 14 focused on pupils' and/or teachers' attitudes and motivations towards astronomy; astronomers; teaching and learning approaches; and science or STEM (science, technology, engineering and mathematics) more broadly. One study explored the impact of a week-long summer programme on secondary school pupils' attitudes towards astronomy. Post-programme surveys indicated that pupils had a greater interest in astronomy subsequent to completing the programme and were more likely to apply to an astronomy degree (Mondim, 2016). Elsewhere, pupils' and teachers' attitudes towards astronomy was seen to increase significantly after engagement in a separate summer programme. Similarly, feedback forms from upper primary/lower secondary school pupils revealed that planetarium experiences improved pupils' interest in science (Görecek Baybars & Çil, 2021).

On a micro level, one study explored the impact of one-off classroom activities on pupils' attitudes towards physics and astronomy, and the value of processes of investigation, autonomous experiences, collaboration with peers, novel experiences and effective differentiation in promoting positive learning experiences in secondary school physics (Bartlett, 2018).

In some instances, studies have delved deeper into the influences of pupils' interests and motivation. Results from the study by (Görecek Baybars & Çil, 2021) demonstrated the value of pupils' opportunity to learn new things and experience learning outside of the classroom. Another study suggested that pupils' willingness to engage in a ubiquitous learning platform (a digital e-learning platform that can be accessed at any time and location) is largely dependent on their perception of its usefulness, ease-of-use and availability of technical support.

Attitudes towards science and scientists, or astronomy and astronomers have also been explored. One study captured written and drawing data from secondary school pupils to distil their mental models of the astronomer. Three models were interpreted from their collected data: a scientific model that highlighted key characteristics of astronomers, a celestial model which offered no human-centred focus, just celestial bodies, and an astronaut model; a drawing of an astronaut in space surrounded by celestial bodies (Görecek Baybars & Kayabaş, 2018). Another focused specifically on attitudes and perspectives among female pupils and the impact of a focused intervention. While at the start, pupils reported a like for STEM subjects, they found them difficult and provided stereotypical representations of scientists. However, after intervention, pupils recognised the wider relevance of science to their lives and their reality (Benitez-Herrera et al., 2019).

Instrument-focus

Three studies in the search focused explicitly on developing or piloting explicit data collection tools to measures attitudes and perspectives of pupils or for evaluating astronomy programmes. Elsewhere, while some studies focused on measuring outcomes of activities or programmes, they also described processes of validating their data collection methods to confirm their effective and reliable use, therefore these studies are also drawn on here.

One study focused specifically on quantitative measures of pupils' attitudes towards astronomy and science in terms of eight key factors: interest in astronomy, interest in science outside of school, practical work in school science, teachers' actions, perceived ability in science, future aspirations, perceived benefits of science and personal relevance of science Bartlett et al., (2018). They describe how these statements can be used pre- and post-programme to measure change in pupils' attitudes.

Several studies used pre-existing data collection instruments that had been used elsewhere. For example, the Astronomy Diagnostic Test (Hufnagel, 2001) created by the Collaboration for Astronomy Education Research which was used in several studies either in full, or part (see Chubko et al., 2019; Kalkan et al., 2014; Rodrigues de Andrade, 2023; Türk & Kalkan, 2015). This is a multiple-choice test that measures pupils' conceptual understanding of introductory astronomy topics including apparent motion of the Sun, scale of the Solar System, Moon phases, seasons, global warming, the nature of light, gravity, starts, cosmology and linear distance scales. Others used an Astronomy Achievement Test (Türk & Kalkan, 2015a), the Test of Astronomy Standards (Slater, 2014), and an Attitudes Towards Astronomy survey (Zeilik et al., 1997).

In terms of studies that described validation processes of their own data collection instruments, there were examples of both quantitative and qualitative method validation. For closed-question surveys, such as Likert-scale statements, several studies employed statistical tests for validation (e.g. Ampartzaki et al., 2024; Ferrari et al., 2024; Kalkan et al., 2014), this provides confidence in the reliability and consistency of their measures which can give others the confidence to adopt these data collection tools. For example, Chen & Lin (2016) used this for their measures of pupils' interest and motivation towards a particular learning platform, and Ampartzaki et al. (2024) used this for their quantitative measures of teachers' astronomy content knowledge. In some qualitative studies, validation and reliability processes included thoroughly piloting data collection instruments, confirming or verifying the evaluator's inferences with participants, or having multiple individuals conduct the analysis independently and then compare their results, where similar results are drawn, this offers a degree of reliability in interpretation of data (Costa et al., 2023; Kaplan, 2014; Kim, 2015).

Focus Summary

- The astronomy education evaluation studies reviewed typically focus on one or more of four key areas: knowledge and understanding; attitude and perspective; experience; and instrument validation.
- Knowledge-focused studies mostly centred on improving pupils' or teachers' astronomy knowledge and understanding or improving teachers' pedagogical content knowledge for teaching astronomy, either in general terms or after engaging with a targeted programme or activity.
- Experience-focused studies evaluated pupils' and/or teachers' experiences of participating and engaging in a particular astronomy education programme or activity. These were not necessarily always about the resulting impact of these programmes, but also the particular experiences of participants during their engagement.
- Attitude-focused studies explored pupils' and/or teachers' attitudes and motivations towards astronomy; astronomers; teaching and learning approaches; and science or STEM (science, technology, engineering and mathematics) more broadly.
- Instrument-focused studies systematically tested and validated the suitability and reliability of particular data collection instruments for capturing subjective data (attitude or experience).

Methods

Broadly speaking, methods for evaluation are either quantitative, qualitative, or a combination of the two, though within these, there are myriad methods, techniques and formats for collecting and recording data. The method of evaluation will depend on a number of factors, primarily relating to the evaluation objectives or research questions, and the data required to answer these. For example, quantitative data involves numerical data and is generally used to test hypotheses and/or measure numerical change, such as astronomy knowledge test scores, enrolments on astronomy programmes, or interest in astronomy measured on a numerical scale. Qualitative methods capture non-numerical data that is typically more detailed and presented as text and/or documents and are generally used when an evaluator seeks to identify how or why something did or did not happen. However, there are other important factors in deciding on a data collection method, for example, your participants and what data they are able to provide; time; resource and opportunity to gather data; and your expertise as an evaluation or researcher.

Of the studies reviewed here, 18 were quantitative studies, 17 were qualitative and 17 were mixed methods. All of the quantitative studies used closed question surveys (for attitude or experience) or tests (for knowledge) to capture data. The majority of quantitative studies explored knowledge gains as a result of an astronomy education activity or programme, either among pupils (Anantasook et al., 2015; Chen & Lin, 2016; de Lima et al., 2018; Dobaria et al., 2022; Ferrari et al., 2024; Kalkan et al., 2014; Kuhn et al., 2017; Türk & Kalkan, 2015b; Zhang et al., 2014) or teachers (Covitt et al., 2015; Setozaki et al., 2018; Stork, 2014). Others however explored key features and characteristics of teacher training programmes across the country (Rodrigues et al., 2024), or quantitative surveys for measuring pupils' attitudes towards astronomy and science (Bartlett, Fitzgerald, Mckinnon, et al., 2018).

Quantitative methods are generally quicker to implement than qualitative methods, for example you can collect 100 survey responses much faster than you can conduct 100 interviews. Their structured nature also makes it relatively straight forward to provide a clear overview of your participants' responses. Some studies highlighted the limitations of quantitative methods, such as the use of closed-questions in terms of the amount and detail of information you can retrieve from participants (Rodrigues et al., 2024), and gives them less autonomy in what information they share. Quantitative methods also generally require larger sample sizes than do qualitative methods.

There appeared to be more diversity in terms of both focus and data collection tools used in the qualitative studies compared to the quantitative studies. There were examples of using open-ended surveys (Aretz et al., 2017; Görecek Baybars & Çil, 2021; Görecek Baybars & Kayabaş, 2018; Kaplan, 2014; Rebull, French, et al., 2018), interviews or focus groups (Benitez-Herrera et al., 2019; Colantonio et al., 2018; Dankenbring & Capobianco, 2016; Isik-Ercan et al., 2014; Kaplan, 2014; Lewis, 2019; Navarro, 2014; Plummer et al., 2015), classroom/activity recordings (de Hosson & Décamp, 2014; Isik-Ercan et al., 2014; Kim, 2015; Plummer, 2015), reflection journals (Lewis, 2019), observations (Isik-Ercan et al., 2014), and drawings (Dankenbring & Capobianco, 2016; Navarro, 2014).

Qualitative methods require fewer participants than quantitative methods as you are providing narrow but deeper insights, they are also much more flexible as they can be adapted throughout the data collection process and give the participant more control over what information (data) they share. However, qualitative methods are more time consuming and can be challenging in terms of providing generalisable conclusions. Several studies reported the limitations of their qualitative studies, most typically in terms of the small sample sizes (Colantonio et al., 2018; Plummer et al., 2015; Rebull, French, et al., 2018).

Mixed method studies can be concurrent or sequential. Concurrent approaches involve implementing multiple methods in parallel, and sequential approaches involve an order, when one method informs or guides another (Creswell & Plano-Clark, 2011). The mixed method studies in this review involved various combinations of data collection tools, often using quantitative methods to explore the 'big picture' and top-level view of their objectives and their participants' knowledge, experience or attitude, and used qualitative methods to explore the detail and reasoning behind these top-level findings. For example, one study found that although pre-service teachers' knowledge test scores improved after a training programme, qualitative data interviews and observations revealed details about teachers' preferences and challenges (Burrows et al., 2016). Alternatively, a study of game-based learning revealed that while pupils' knowledge increased, their reflections in interviews demonstrated a lack of confidence in their full understanding, and teachers commented their view that game-based learning worked best as a revision tool to reinforce learning rather than a tool for learning new topics (Cardinot & Fairfield, 2019b).

Methods Summary

- Evaluation methods can employ quantitative (numerical), qualitative (non-numerical) or mixed methods for data collection.
- Quantitative methods typically use closed-question attitude or experience surveys, or knowledge-based tests. They are generally used to test hypotheses and/or measure numerical change.
- Quantitative methods are quick to implement and obtain data and their structured nature make it easy to summarise your data. However, they are rigid, limiting the richness of the data that can be collected and generally require large sample sizes.
- Qualitative methods capture more detailed data, most often represented as text and/or documents. They are useful for identify why something did or did not happen, or did or did not work. However, they are time and resource heavy and are limited by their typically small sample size.
- Mixed methods involve implementing both quantitative and qualitative approaches either concurrently or sequentially.

Study Designs

The retrieved studies also employed a number of different designs for implementing data collection methods for evaluation. Four key designs were employed to evaluate astronomy education programmes, practices and circumstances: cross-sectional studies, control group comparisons, repeated measures, and case studies.

Cross-Sectional

Cross-sectional designs capture data from a group or audience at one point in time. They are well suited to assessing the status of group's understanding or attitude. For example, one study sought to explore key characteristics of astronomy teacher training programmes in Chile by collecting survey data from programme leads. Their results highlighted the geographical spread of programmes, delivery formats and durations, content, and curriculum alignment (Rodrigues et al., 2024). In another study focusing on teachers, cross-sectional survey data revealed a lack of astronomy content knowledge among primary and secondary school teachers and key differences in terms of the level of teacher training received. In the context of pupils, a study explored secondary school pupils' understanding of celestial motion through a social and cultural context. Cross-sectional data revealed a connection between understanding and social and cultural contexts such as positions of chapels, sanctuaries and modern buildings in relation to the Sun's path across the sky, and the influence of celestial motion on daily lives.

However, some studies will use cross-sectional data to assess impact by gathering data after an intervention, programme or activity has occurred and asking participants for retrospective feedback. For example, asking participants if they feel their perception of careers in astronomy has changed after the activity, or how confident they feel in carrying out image processing compared to before a training programme. In one study, secondary school pupils reported liking astronomy more than before after engaging in a summer programme (Mondim, 2016). Cross-sectional studies generally require less time as you only need to implement data collection once but can be difficult in demonstrating change.

Repeated Measures

In contrast to cross-sectional studies, a repeated measures design captures data from the same group over multiple timepoints, most typically, before and after implementation of an astronomy education activity or programme, and sometimes, a follow-up several weeks, months or a year later. They often involve repeating the same data collection measure to permit direct comparisons across timepoints.

Several repeated measures studies demonstrated an increase in knowledge or improvements in attitudes subsequent to engagement in astronomy education activities. Whether that be short-term engagements through game-based learning (Cardinot & Fairfield, 2019a), or multisensory workshops (De Leo-Winkler et al., 2019), or longer-term summer projects (Türk et al., 2015). One study implemented the Astronomy Diagnostic Test with secondary school pupils and teachers before and after a five-day science school. Having data before and after implementation, and from two participant groups permitted investigation of changes both within and between groups (Kalkan et al., 2014). Their results revealed that pupils and teachers generally had the same level of content knowledge at the baseline and same extent of improvements after the science school.

Other studies have employed a longitudinal design to capture participants views or understanding both immediately after an intervention or experience and again several weeks or months later to explore whether any outcomes were sustained. Zimmerman et al. (2014) investigated different technology-enhanced learning environments on pupils' astronomy knowledge retention. Their design involved measuring two groups of secondary school pupils' knowledge immediately and six weeks after engaging in a planetarium show delivered either within a dome environment or on a computer. They found a significant increase in test scores through both mediums, however long-term retention was more favourable among pupils who experienced the dome. This more favourable approach for long-term retention would not have been known had they only measured pupils' knowledge immediately after implementation.

Control Groups

Where repeated measures design permit comparisons of the same group of participants at different time points, control and experimental/treatment group designs permit the comparison of different treatments or experiences. In the literature, control groups were used to evaluate the impact of an astronomy education implementation or intervention, for example a particular teaching method, an educational resource, or a programme. Such studies involved a control group who experienced their usual day-to-day teaching and learning approaches, and a treatment or experimental group that experienced the astronomy education intervention. Both groups participated in the same data collection methods which permitted direct comparisons of results. Control group designs are often regarded as providing some of the more robust evaluation data as they offer confidence that any changes observed among the experimental group but not the control group, are a direct result of the intervention.

One study used a control group design to explore how engaging in scientific inquiry within a multivariable context can promote understanding among secondary school pupils as measured by knowledge test scores. Their results indicated that the intervention group significantly outperformed the control group on the knowledge test in terms of their investigation and argumentation skills (Kuhn et al., 2017). Elsewhere, a similar design was implemented to evaluate pupils' understanding of the notions of velocity and of gravitational attraction and inertia when experiencing 'enacted astronomy'; 'acting out' the motions of the Solar System with pupils moving in planetary orbits, compared to traditional instruction (Rollinde, 2019). Pupils who experienced instruction in a planetarium environment significantly improved in their knowledge test scores and misconceptions, whereas no changes were seen for the control group who experienced their usual classroom environment (Türk & Kalkan, 2015b).

Some studies were seen to combine both control groups and repeated measures designs. Here, both the control and experimental group completed pre- and post-data collection. Such a design permits not only investigation of changes to participants experiencing the astronomy education programme or activity, but a benchmark of comparison. This repeated measures with a control group adds greater confidence that any changes to the treatment group that did not happen to the control group are down to the influence of the intervention. For example, Dankenbring & Capobianco (2016) used both a control group and repeated measures to explore the impact of engineering-design tasks on pupils' conceptual understanding of Sun-Earth relationships compared to traditional teaching methods. Similarly, a study on the value of AR in astronomy teaching in secondary schools revealed that while the control group did not significantly improve in knowledge-test scores, the treatment group improved significantly with large effect sizes (Ferrari et al., 2024).

Case Studies

Case study designs involve an in-depth study of a particular 'case' or cases. In an educational setting a case could be a particular class of pupils, a particular school or even individual teachers or pupils. Due to their in-depth nature, case studies often involve multiple methods of data collection and involve qualitative methods. Case studies are very good at studying a specific context in great detail but can be limited in how much their findings can be transferable to other settings.

In the reviewed literature, case studies were typically used to explore teachers' and/or pupils' experiences of a particular astronomy education course or programme, as well as the impact on their knowledge. One study collected data through course assignments, focus groups, lesson plans and reflection journals to explore how experiencing primary school teaching prepares primary school teacher candidates for practice. Their findings urged the value of training teachers through processes of inquiry rather than teaching them about inquiry, and the value of physical investigations, questioning and scientific modelling in developing science skills (Lewis, 2019). Elsewhere, a case study design with repeated measures captured data from both pupils and teachers to explore how 3D visualisations support primary school pupils' astronomy learning. The study employed pre- and post- interviews with pupils and teachers, audio recoded lessons, and classroom observations which revealed greater understanding among pupils after using 3D software and teachers observed an increased motivation among their pupils but recognised variation between individuals (Isik-Ercan et al., 2014).

Study Designs Summary

- The main study designs employed in the reviewed literature were cross-sectional, repeated measures, control groups and case studies.
- Cross-sectional designs capture data from an audience at one point in time and so are well suited to diagnostic evaluation, or for collecting information about the influence of an astronomy education programme or activity retrospectively.
- Repeated measures involve capturing data from the same group multiple times, thus permitting exploration of how participants' data, such as knowledge or attitude, changes over time, for example, after experiencing a particular programme or teaching method
- Control groups involve collecting the same data from multiple groups who have had different treatments or experiences, for example traditional classroom teaching compared to technology-based teaching.
- Case studies are an in-depth study of a particular 'case', for example a particular class of pupils. They provide a detailed but narrow evaluation.

GOOD PRACTICES

While the 52 evaluation studies in this review were not formally quality assessed, a number of good practices were identified, as well as some limitations. Astronomy educators may want to consider these when planning their evaluation studies and are summarised and described in the following subsections.

Aligning Evaluation Measures with Programme Objectives

Evaluation is done best when it goes beyond simple feedback questions about participants' enjoyment of an activity or programme, or what they liked best and least. Instead, the questions and methods of evaluation should align with what impact you are aiming to have on your audience or the information you need to acquire. Ultimately, evaluation is about how well something works and the extent to which the intended outcomes were achieved. In order to identify these, you should first consider what are your indicators of success, and then explore what evaluation methods, designs and data collection techniques will provide you with this information.

Triangulation

Triangulation involves collecting data either from multiple sources (e.g. pupils, teachers and parents), multiple methods (e.g. surveys and interviews), or multiple evaluators. Triangulation is a common method for increasing the reliability of your evaluation results. Having data from multiple sources provides a more holistic perspective of your programme or activity and understanding of different perspectives. Using multiple data collection techniques permits a broader and often more detailed study. Having more than one evaluator or research review and analyse the results provides greater confidence in interpretations of data and in drawing inferences and conclusions. Generally speaking, triangulation allows you to check for consistencies or discrepancies among your data and overcome potential biases in interpretation.

Establish a Baseline or Benchmark

If you are using evaluation to identify the impact or outcome of an astronomy education activity or programme, having a baseline or benchmark to compare your results with is a valuable way to provide confidence in your findings. Establishing a baseline involves understanding the status of your participants' before they have engaged with your activity or intervention so you can compare this directly with results after their engagement. For example, implementing the same knowledge test before and after an astronomy course. To strengthen your results even more, you could also include a control group who provide data at the same point as your participants but who do not engage in your astronomy education activity. If those that engage in your activity improve but the control group doesn't, then you can be more confident of the influence of your activity. However, if both groups improve, then you can assume the change was a result of something else other than your activity.

If you are running a programme with multiple groups or cohorts, for example, repeating implementation with different groups each year you may want to also repeat your evaluation to track your results over time. This can help you identify any patterns or trends in your data, or perhaps areas that you need to focus on.

McLin & Cominsky (2014) did this in their evaluation of the NASA Astrophysics Educator Ambassador Program.

Engage with Stakeholders

There are many groups and individuals involved in education of young people, so you may want to consider including some of these in your evaluation either as participants who provide data and feedback, or as collaborators in designing your evaluation protocol. This could include teachers and parents, but also education policy makers, education ministers, examination boards, local authorities and many others. Who you work with in your evaluation will depend on the objectives you are trying to achieve.

For example, if you want to inspire young people's interest in astronomy through an astronomy summer camp, you may want to ask parents if they feel their child talked about astronomy more after the camp, or showed an enthusiasm. If you are trying to explore how astronomy can be embedded in the curriculum, you may want to engage with policymakers.

Consider Contextual or Cultural Factors

Evaluation isn't typically used to yield generalisable findings that can be applied in all context and settings. What works in one classroom or school system, may not work elsewhere. For example, the studies reviewed here that involved aboriginal and indigenous astronomy content (Ruddell et al., 2016; Wyatt et al., 2014) is unlikely to yield such positive outcomes if it were implemented in a Western country. Is it therefore important that when disseminating the findings of your evaluation, you describe the surrounding context, such as cultural, socioeconomic and environmental factors. This way, other astronomy educators can consider whether the findings from your study are likely to withhold in the context of their audiences.

It is also important to consider the evaluation methods and data collection techniques you use and whether they are appropriate for your audience, in terms of language, cultural norms and special/additional educational needs. In the studies by (De Leo-Winkler et al., 2019; Guiotto et al., 2023) that involved partially sighted and deaf pupils, they will have had to consider how they gather data when traditional survey or interview methods would not have been suitable.

Analyse Data Appropriately

Methods of analysing your data will depend on the type of data you have (i.e. quantitative or qualitative) and the format of it. It is important to consider what kind of analyses you need or want to do whilst planning or designing your data collection methods. For example, you may ask participants the question, 'how much did you learn about stars today?', they could respond to this in a number of ways, it could be an open (qualitative) question, or they could be given multiple options to choose from, for example a scale of 0 to 10, or just descriptors like 'nothing', 'not much', 'some', or 'a lot'. You should consider what analyses you need to do and which of these question styles will give you the most helpful data in answering your evaluation objectives.

There are various statistical analysis methods for quantitative data, but they will need to meet certain requirements so it's important to be aware of these before you collect evaluation data. Qualitative data can be analysed in different ways such as through thematic analysis or discourse analysis, again you should consider what approach is appropriate for your data and your evaluation objectives. If you triangulate multiple methods or sources, it is important to consider if and how you will combine your data and how you will approach analyses of these different data.

Ethical Protocols

Ethical considerations and protocols are important when collecting 'human data'. Human data describes any data that is collected about or from people, whether identifiable or anonymous. In general terms, ethical considerations, in the context of research and evaluation, are concerned with what is considered to be morally right or wrong and ensure that evaluation or research is conducted fairly and respectfully. This largely focuses on, though is not limited to, how you engage with your 'participants' (the people participating in your evaluation/research). Ethical considerations include factors such as ensuring you keep participants' information private, minimising risks to them, being clear and transparent about why you are asking them to participate and what you will do with their information. Ethical protocols are also about balancing the demands on your participants and ensuring the evaluation does not detract from the more important astronomy education activities, or you are not asking too much of them. For example, you wouldn't interview a teacher for two hours to get their feedback on a thirty-minute teaching training event. Not all of the retrieved studies in the review included a description of ethical considerations, protocols or review by a relevant committee, however, that is not to say they did not acquire this. Academic journals are increasingly requesting evidence of ethical protocols as part of their publication requirements, and generally, it is encouraged as best practice when collecting data from human participants.

If you are conducting an evaluation of any astronomy education activity, project, initiative or implementation and you intend to make your results or findings available to a wider audience beyond your evaluation team, we encourage you to acquire an ethics review from a relevant committee. Ethics committees are established groups of experienced and/or or expert individuals in research practice (usually in the same or a relevant field) that review evaluation and research protocols involving involve human participants or human data. This review ensures that individuals' rights, safety, dignity and wellbeing are protected and not damaged because of the research. individuals here include not just the participants, but also you as the evaluator.

USING EVALUATION RESULTS

Evaluation findings and results can be useful in a variety of ways, from ongoing development of astronomy education material or programmes to informing policy. The following sections describe how you can leverage the outcomes of your evaluation studies.

Continuous Improvement

Evaluation practices are critical when you are looking to continuously improve or develop an astronomy education programme or activity. In the earlier section 'Types of Evaluation', diagnostic, formative and summative evaluations were described. Formative evaluation, in particular, is invaluable in supporting ongoing adaptation of teaching and learning methods, enabling educators to refine and enhance activities over time. Evaluation can inform what you do in the future, how you develop your current programmes and what additional projects or activities might be of benefit to your audience. This ensures your astronomy education activities are dynamic, responsive and impactful to your audience.

Data Informed Decision Making

Evaluation is all about gathering evidence which can offer valuable insights to inform decision making around instruction, practice, policy and curriculum design. Evaluation results provide evidence to different stakeholders that enables decisions on what programmes to continue, upscale, or even discontinue if they are demonstrating little success. It can also help to refine programmes or materials for different audiences, or even help you decide which audiences you want to focus on in the future, e.g. primary or secondary school pupils, teachers, pupils with additional learning needs.

In some instances, evaluation outcomes may be able to guide decisions about curriculum changes, for example evidence of high interest and motivation in astronomical topics could indicate a benefit of greater focus on astronomy content at particular states of education to promote pupil engagement. Evaluation of different teaching and learning methods such as project-based learning, technology-assisted learning, or hands-on learning could reveal implications for pedagogical approaches, and which should be encouraged or discouraged in classrooms.

Ultimately, the more evidence you gather from evaluations of astronomy education resources or programmes, the more understanding you can acquire in how best to allocate resources and meet the needs of your audiences.

Supporting Professional Development

Evaluating astronomy education practices is valuable in supporting the professional development of educators, not only yielding greater understanding for you as the evaluator, but to them as educators. This could be in the context of using astronomy as a point of engagement for other curriculum content or for specifically teaching astronomy as its own subject. By identifying effective pedagogical approaches and teaching methods, evaluation provides valuable insights into what works well in different classroom contexts. They can uncover gaps in teachers' knowledge or limitations in instructional practices (e.g. Kaplan, 2014; Rodrigues de Andrade, 2023; Setozaki et al., 2018), or what enhances pupils' engagement with scientific topics (Cardinot & Fairfield, 2019a; De Leo-Winkler et al., 2019; Lee & Feldman, 2015), this understanding offers an opportunity to design targeted training programmes for educators. Such professional development can help ensure that educators are better equipped to teach astronomy or related topics effectively, fostering both their own development and the success of their pupils.

Funding Applications

Findings from evaluations can provide compelling evidence for applications for funding. By providing evidence of impact through the evaluation of pilot activities or small-scale projects, educators can provide justifications for continuation or upscaling of these initiatives into larger-scale or longer-term programmes. Not only does such information demonstrate possible pathways of impact, but it fosters confidence in the funder that your team or organisation is able to deliver successful outcomes. Thorough evaluation that is well-articulated can be a key strength of funding proposals, increasing the likelihood of success. Furthermore, embedding an evaluation plan for a funding proposal will also be well received as if granted funding, funders will want to see the subsequential impact of implementation.

CONCLUSION

Evaluation is a powerful tool that can reveal important outcomes and impact and drive meaningful change in astronomy education. By gathering and analysing data, astronomy education practitioners are able to systematically assess the effectiveness of their programmes and activities, identify areas for improvement and make informed decisions to enhance teaching and learning. Whether evaluating pupil outcomes, teacher training, or curriculum design, evaluation is essential to fostering a culture of accountability, equity and continuous development.

This document aims to provide a foundational understanding of evaluation, covering key concepts, types of evaluation, and crucially, specific examples of different methods, focus areas and designs employed in astronomy education evaluation studies. Good practices have been outlined that can facilitate effective, ethical and impactful evaluations, as well as an emphasis on how to translate evaluation results into actionable strategies that benefit not only pupils and educators, but wider stakeholders.

Evaluation should not be regarded as a one-off activity, but an ongoing process, embedded in astronomy education practices and that adapts to evolving goals, objectives and audience need. When done well, it has the potential to improve not only individual programmes or activities, but broader education systems or practices. Application of the principles and practices shared in this guide can help towards fostering an informed, responsive and effective network of astronomy educators.

The following section provides more detail of the 52 studies encompassed in this evaluation review and the relevant references for further reading. These may be helpful in identifying suitable examples of protocols, methods, and analyses when designing astronomy education evaluation studies.

REVIEWED LITERATURE: SUMMARY

Title	Authors	Focus	Evaluation / Research Objectives	Participants	Design	Methods
A Characterization of Astronomy Teacher Professional Development Programs in Chile	(Rodrigues et al., 2024)	Status	Explore key characteristics of teacher professional development programmes in Chile	Primary and secondary school teachers	Cross-sectional	Quantitative: Closed-question survey
A Scientific Modelling Sequence for Teaching Earth Seasons	(Covitt et al., 2015)	Knowledge	To assess the instructional effectiveness of an Earth seasons instruction model	Secondary school teachers	Repeated measures	Quantitative: Closed-question knowledge test
A Week-long Summer Programme in Astronomy for High-school Students	(Mondim, 2016)	Attitude	To evaluate the effectiveness of a summer programme on uptake and interest in astronomy	Secondary school pupils	Cross-sectional	Quantitative: Closed-question survey
Assessing the Initial Outcomes of a Blended Learning Course for Teachers Facilitating Astronomy Activities for Young Children	(Ampartzaki et al., 2024)	Knowledge	1. What are the primary outcomes of an initial assessment of a blended learning course for teachers? 2. What is the impact of the programme on the content knowledge and pedagogical content knowledge of teachers? 3. What aspects of the programme were most challenging for teachers? 4. What improvements were identified as necessary?	Primary school teachers	Repeated measures	Mixed methods: Closed-question survey, interviews
Astronomy and Science Student Attitudes (ASSA): A Short Review and Validation of a New Instrument	(Bartlett, Fitzgerald, McKinnon, et al., 2018)	Attitude, Tool	To develop a survey instrument that reliably and validly measures the effectiveness of relevant intervention projects in astronomy education	Secondary school pupils	Cross-sectional	Quantitative: Closed-question survey
Authentic Science Experiences: Pre-Collegiate Science Educators Successes and Challenges During Professional Development	(Burrows et al., 2016)	Knowledge, Experience	1. How do pre-collegiate STEM educators describe successes and challenges when immersed in astronomy authentic science practices? 2. What are the educators' perceptions before and after astronomy authentic science practices?	Primary and secondary school teachers	Repeated measures	Mixed methods: Interviews, journalling, observations, closed-question survey, closed-question knowledge test

			3. How do pre-collegiate STEM educators show learning gains on pre to post content tests during the astronomy authentic science practices?			
Basic Astronomy Concepts in The Footsteps of Eratosthenes	(Kalkan et al., 2014)	Knowledge	To determine change in primary school pupils' and teachers' knowledge of basic astronomy concepts	Secondary school pupils and secondary school teachers	Repeated measures	Quantitative: Closed-question knowledge test
Can Engaging in Science Practices Promote Deep Understanding of Them?	(Kuhn et al., 2017)	Knowledge	What are students' understanding of science as a practice entailing the debate of alternative claims in a context of evidence?	Secondary school pupils	Control group	Quantitative: Closed-question knowledge test
Comparison of student learning about space in immersive and computer environments	(Zimmerman et al., 2014)	Knowledge	1. How does student learning about space compare in digital and computer environments immediately and six weeks after. 2. To improve students' desirability for a career in space	Secondary school pupils	Repeated measures	Mixed methods: Closed-question knowledge test, attitude survey, focus groups
Design and development of a learning progression about stellar structure and evolution	(Colantonio et al., 2018)	Knowledge	How do students' progress their understanding about stellar structure and evolution when exposed to iteratively designed instructional activities in a learning progression framework?	Secondary school pupils	Repeated measures	Qualitative: Interviews
Designing Science Learning with Game-Based Approaches	(Liu et al., 2014)	Knowledge	What is the effect of a game-like environment on sixth-grade students' science learning and what do they learn by using the environment?	Primary school pupils	Repeated measures	Mixed methods: Closed- and open-question knowledge test
Determination of Secondary School 4th Grade Students' Mental Models of the Astronomer	(Görecek Baybars & Kayabaş, 2018)	Attitude	To determine secondary school 4^{th} grade students' mental models of the astronomer	Secondary school pupils	Case study	Qualitative: Open-question survey
Development and Evaluation of a Construct Map for the Understanding of the Expansion of the Universe	(Aretz et al., 2017)	Knowledge	Test an existing structural setup of student understanding of the expansion of the universe: 1. Is it possible to assign students to just one or two adjacent levels on the basis of a further developed and potentially adjusted construct map? 2. Is there a correlation between the misconceptions of the universe having an edge and/or centre and the concept of an expanding space?	Secondary school pupils	Cross-sectional	Qualitative: Open-question survey

Development and evaluation of a context-aware ubiquitous learning environment for astronomy education	(Chen & Lin, 2016)	Knowledge, attitude	1. Improve student interest and motivation in learning astronomy 2. Assess student willingness to use 'context aware astronomy learning systems'	Secondary school pupils	Control group, repeated measures	Quantitative: Closed-question knowledge test
Development of a Learning Progression for the Formation of the Solar System	(Plummer et al., 2015)	Knowledge	How can we describe levels of sophistication in student thinking about the Solar System and its formation using a 'learning progression' framework?	Secondary school pupils	Cross-sectional	Qualitative: Interviews
Dreamtime Astronomy: development of a new Indigenous program at Sydney Observatory	(Wyatt et al., 2014)	Experience	How can Aboriginal content be relevant and memorable to students who typically have never been exposed to Aboriginal astronomy prior to a visit to Sydney Observatory?	Secondary school teachers	Cross-sectional	Teacher feedback (detail on methods not specified)
Edukoi: developing an interactive sonification tool for astronomy between entertainment and education	(Guiotto et al., 2023)	Experience	1. Do less-sighted and BVI children understand that sonification makes astronomy accessible? 2. Do sonifications make learning science more engaging for everyone? 3. Are users of Edukoi able to recognise shapes, recognise colours, and explore and understand the content of astronomical images?	Middle school pupils	Repeated measures	Qualitative: Interviews
Empowering Prospective Teachers to Become Active Sense-Makers: Multimodal Modelling of the Seasons	(Kim, 2015)	Knowledge, experience	How do Singaporean prospective teachers construct and use multimodal models to develop a deeper understanding of the seasons when co-designing an Embodied Modelling-Mediated Activity?	Pre-service primary school teachers	Case study	Qualitative: Video recording of implementation, artifacts, drawings, screen captures, lesson plans and presentation slides.
Evolutionary Maps: A new model for the analysis of conceptual development, with application to the diurnal cycle	(Navarro, 2014)	Knowledge, tool	Are evolutionary maps a methodology that accurately maps the web of paths that children may follow in the development of concrete concepts?	Primary school pupils	Control group	Qualitative: Interviews involving open-questions and drawing tasks
Examining Elementary School Students' Mental Models of Sun-Earth Relationships as a Result of Engaging in Engineering Design	(Dankenbring & Capobianco, 2016)	Knowledge	1. Determine students' conceptual understanding of Sun-Earth relationships. 2. Compare conceptual understanding of students who completed an engineering	Primary school pupils	Control group, repeated measures	Mixed methods: Closed-question knowledge test, draw-and -explain activity, interviews.

			design to students who completed traditional science learning activities?			
Game-Based Learning to Engage Students with Physics and Astronomy Using a Board Game	(Cardinot & Fairfield, 2019a)	Knowledge, experience, attitudes	1. Does the use of GBL influence student learning of astronomy topics? 2. What are student and teacher perceptions of learning through games?	Secondary school pupils, secondary school teachers	Repeated measures	Mixed methods: Feedback surveys, observations, focus groups, open- and closed-questions knowledge test.
Improving Astronomy Achievement and Attitude through Astronomy Summer Project: A Design, Implementation and Assessment	(Türk et al., 2015)	Knowledge, attitudes	1. What is the impact of the project on the participants' astronomy achievement and their attitude to astronomy? 2. What is the relationship between the students, preservice teachers and in-service teachers' astronomy achievement and their attitudes to astronomy?	Secondary school pupils, secondary school teachers	Repeated measures	Quantitative: pre and post closed-questions knowledge tests, pre and post closed-questions survey.
Indigenous Sky Stories: Reframing How we Introduce Primary School Students to Astronomy: a Type II Case Study of Implementation	(Ruddell et al., 2016)	Experience	Assess the impact of the indigenous sky stories programme on students	Primary school pupils	Repeated measures, case study	Mixed methods: Group interviews, observations of engagement, photographs, administrative attendance data.
Learning Science Through Enacted Astronomy	(Rollinde, 2019)	Knowledge, attitudes	Assess the existence of conceptions and to evaluation the level of understanding of two fundamental notions among 14-16 year old pupils	Secondary school pupils	Control group	Mixed methods: Open- and closed-question knowledge test, interviews.
Major outcomes of an authentic astronomy research experience professional development program: An analysis of 8 years of data from a teacher research program	(Rebull, French, et al., 2018)	Knowledge, experience	How do educator participants describe the major changes and outcomes in themselves fostered by the project?	Primary and secondary school teachers	Repeated measures, case study	Qualitative: Open-question survey
Motivations of educators for participating in an authentic astronomy research experience professional development program	(Rebull, Roberts, et al., 2018)	Attitude, experience	How do participating teachers describe their motivations for participating in NITARP? (NASA/IPAC Teacher Archive Research Program)	Secondary school teachers	Repeated measures	Qualitative: Open-question feedback forms

Photographs and Classroom Response Systems in Middle School Astronomy Classes	(Lee & Feldman, 2015)	Knowledge, experience	1. To investigate students' learning experiences when classroom response systems are combined with visual stimuli. 2. What is the pedagogical role of photographs in a discussion-oriented pedagogy?	Middle school pupils, middle school teachers	Repeated measures	Mixed methods: Closed-question knowledge test and survey, daily feedback, interviews.
Practice what you teach: How experiencing elementary school science teaching practices helps prepare teacher candidates	(Lewis, 2019)	Knowledge, experience	1. What are teacher candidates' experiences and evaluations of the science methods course aimed at merging the teaching of science content with the teaching/modelling of inquiry-based instruction? 2. Which aspects of the curriculum do teacher candidates say help them learn to science content and pedagogical content knowledge?	Pre-service primary school teachers	Case study	Qualitative: Course assignments, focus groups, lesson plans, researcher reflection journals.
Research on the Pseudo-scientific Beliefs of Pre-service Science Teachers: a Sample From Astronomy-Astrology	(Kaplan, 2014)	Attitude, knowledge	1. What are pre-service science teachers' perceptions on astrology? 2. What are pre-service teachers' abilities to distinguish between science and pseudo-science through examples of astronomy-astrology?	Pre-service primary school teachers	Cross-sectional	Qualitative: Open-question knowledge test, open-question survey, interviews
Science communication through astronomy education: The creation, implementation, and assessment of Porto Planetarium's science education strategy	(Costa et al., 2023)	Experience	What is the effectiveness of a planetarium education programme in connecting science education and science communication?	Primary and secondary school teachers	Cross-sectional	Mixed methods: Open- and closed-questions survey, observations, interviews
SOLO taxonomy as EFL students' disciplinary literacy evaluation tool in technology-enhanced integrated astronomy course	(Chubko et al., 2019)	Knowledge	How does an integrated astronomy course enhanced with digital storytelling affect English-as-foreign-language students' disciplinary literacy development in the astronomy context?	Secondary school pupils	Repeated measures	Quantitative: Closed-question knowledge test
Students' Scientific Evaluations of Astronomical Origins	(Dobaria et al., 2022)	Knowledge	1. What are the levels of students' evaluations when they engage in two instructional scaffolds on astronomy topics?	Secondary school pupils, pre-service secondary school teachers	Repeated measures	Quantitative: Explanation tasks, plausibility ratings, closed-questions knowledge test

			2. How do students' plausibility judgements and knowledge about astronomy topics change over the course of these two instructional scaffolds? 3. How do relations between students' evaluations, plausibility judgements, and knowledge compare between the two scaffolds?			
Thai Students Understanding about Celestial Motion within Their Social and Cultural Context	(Anantasook et al., 2015)	Knowledge	1. Give empirical data on students' perspective on astronomy teaching 2. What are the implications for the development and evaluation of teaching using a sociocultural approach in the celestial motion unit for Thai students?	Secondary school pupils	Cross-sectional	Quantitative: Closed-question knowledge test
The Development and Evaluation of an Augmented Reality-based Armillary Sphere for Astronomical Observation Instruction	(Zhang et al., 2014)	Knowledge, attitude	1. Determine whether improving environmental factors and the learning tool is effective in improving cognitive factors 2. Determine whether changing learning tools enhances performance factors 3. Determine whether an improvement in cognitive and performance factors improves affective factors 4. Determine whether an improvement in cognitive and performance factors prolongs the effects of affective factors	Primary school pupils	Control group, repeated measures	Quantitative: Closed-question knowledge test and survey
The Effect of Planetariums on Teaching Specific Astronomy Concepts	(Türk & Kalkan, 2015b)	Knowledge	What changes occur in academic achievements and misconceptions related to specific astronomy concepts after the instruction of students in a planetarium environment compared with those in education classroom environment?	Secondary school pupils	Control group, repeated measures	Quantitative: Closed-question knowledge test
The Vibrating Universe: Astronomy for the Deaf	(De Leo-Winkler et al., 2019)	Experience	Evaluate the workshop's effectiveness in providing a positive, engaging and fun experience in terms of: 1. Students' reactions to 19 different sounds 2. Students' overall least/favourite vibration 3. Suitability of the duration of activities	Primary and middle school deaf pupils	Cross-sectional	Quantitative: Closed-question feedback form

Unlocking the cosmos: evaluating the efficacy of augmented reality (AR) in secondary education astronomy instruction	(Ferrari et al., 2024)	Knowledge	1. Is there a significant difference in astronomy literacy levels between pre and post averages of control group? 2. Is there a significant difference in astronomy literacy levels between pre and post averages of experimental group who were exposed to AR-based instructional interventions? 3. What is the magnitude of the impact of AR on astronomy literacy?	Secondary school pupils	Repeated measures, control groups	Quantitative: Closed-question knowledge test
Use of portable planetariums in teaching of astronomy: an activity example	(Görecek Baybars & Çil, 2021)	Experience, attitude	Understanding pupils' experiences of planetarium activities	Upper primary/lower secondary school pupils	Cross-sectional	Qualitative: Open-question feedback form
Using Ancient Chinese and Greek Astronomical Data: A Training Sequence in Elementary Astronomy for Pre-Service Primary School Teachers	(de Hosson & Décamp, 2014)	Knowledge	1. To understand the feasibility and acceptability of the training session 2. Evaluate materials of session for other training sessions	Pre-service primary school teachers	Case study	Qualitative: Recordings and transcriptions of training implementation, lesson materials, lesson guidance notes
We put on the glasses and Moon comes closer! Urban Second Graders Exploring the Earth, the Sun and Moon Through 3D Technologies in a Science and Literacy Unit	(Isik-Ercan et al., 2014)	Knowledge, experience	1. How does a science and literacy unit that employs multiple techniques including 3D visualisation support urban second graders' astronomy learning? 2. How do classroom teachers and children perceive the use of 3D in school learning?	Primary school pupils, primary school teachers	Case study, repeated measures	Qualitative: Interviews, audio recorded and transcribed lessons, classroom observations
What's That Object? Learning Astronomical Concepts Through the Use of The Aladin Program And Manipulation Of Astronomical Images	(de Lima et al., 2018)	Knowledge	Evaluation of the implementation of the activity	Primary school pupils, primary school teacher	Repeated measures	Quantitative: Closed-question knowledge test
A Holistic View of Using Real and Virtual Models in Teaching Astronomy Concepts	(Tsihouridis et al., n.d.)	Knowledge, experience	To test the effectiveness of two teaching models (self-constructed real model and a 3D virtual model) in terms of learning Earth-Moon-Sun movements: 1. To what extent and depth can the use of these models detect students' alternative ideas focusing on the degree of their persistence?	Secondary school pupils	Repeated measures	Mixed methods: Open- and closed-question knowledge test, observations, interviews

			2. To what extent do the results of these detections differ from the results of the administered questionnaire? 3. Which are the learning outcomes of using these models as didactic tools for teaching astronomy concepts? 4. To what extent can the use of these models help students meta-cognitively regarding astronomy concepts?			
An Evaluation of Secondary School Students' Learning Experiences with Astronomy-based Physics Outreach Activities	(Bartlett, 2018)	Attitude, experience	Investigate secondary school students' perceptions of their learning experiences in physics at school and during the implementation of educational activities developed around an astronomy context: 1. How are students' perceptions of positive learning experiences promoted within a physics lesson? 2. How do teachers' planning and delivery of the activities influence students' learning experiences? 3. How does the implementation of the activities reveal the opportunities and limitations they lend to students' learning experiences? 4. What are the implications for the future development and utilisation of physics activities for developers and teachers?	Secondary school pupils, secondary school teachers	Repeated measures, case study	Mixed methods: Closed-question surveys, observations, interviews
Astronomy School Education in Chile: Perspectives and Connections from the National Curriculum, Teachers' Content Knowledge, And Professional Development Programs	(Rodrigues de Andrade, 2023)	Knowledge, status	1. Identify astronomy content in the Chilean curriculum 2. Map primary and secondary school teachers' astronomy knowledge 3. Evaluate the effectiveness of teacher training programmes	Primary and secondary school teachers	Repeated measures	Mixed methods: Document analysis, closed-question survey, open-question survey
Contemporary discipline-based astronomy education research study of K-12 teachers'	(Stork, 2014)	Knowledge	Examine the existing knowledge state of teachers tasked with teaching basic astronomy concepts dictated by national science education reform documents	Primary and secondary school teachers	Cross-sectional	Quantitative: Closed-question knowledge test

astronomy knowledge using the test of astronomy standards						
Daily life astronomy activity and its contribution to children and educators in science education	(Tomita, 2019)	Knowledge, attitudes	Identify the influence of daily life astronomy activity in nurseries and after-school care in improving children's and educators' scientific views and skills	Pre-school pupils and pre-school teachers	Cross-sectional	Qualitative: Observations, open-question survey
Methods of Engaging Preschool-age Children in Science Practices During Astronomy Activities	(Plummer, 2015)	Knowledge	Illustrate ways in which children engaged in the practices of science through a series of 12 museum workshops	Pre-school pupils	Cross-sectional	Qualitative: Video recordings of workshops
NASA Astrophysics Educator Ambassador Program	(McLin & Cominsky, 2014)	Experience	Evaluate the success and impact of the NASA Astrophysics Educator Ambassador Program	Primary and secondary school teachers	Cross-sectional (but yearly implementation with different cohorts)	Mixed methods: Closed- and open-question feedback survey
Pursuing gender equality in Astronomy in basic education: the case of the project "Girls in the Museum of Astronomy and Related Sciences"	(Benitez-Herrera et al., 2019)	Experience, attitude	Evaluate phase 1 of 'Girls in the Museum' programme to verify possible changes in students' viewpoints and perceptions towards science, in particular, astronomy	Secondary school pupils (females)	Repeated measures	Qualitative: Interviews
Scoping the window to the universe; Design considerations and expert evaluation of an augmented reality mobile application for astronomy education	(Antoniou et al., 2018)	Knowledge, experience	Evaluate the value of a card-based AR mobile app that allows learners to tacitly interact with Solar System objects	Primary school pupils, primary school teachers	Control group	Mixed methods: Closed- and open-question knowledge tests, closed-question survey, open-question survey
The NASA Galileo Educator Network: Using Astronomy to Engage Teachers in Science Practices	(Kruse et al., 2015)	Experience	Gauge effectiveness of the NASA Galileo Educator Network in terms of impact and implementation	Teacher trainers	Repeated measures	Mixed methods: Closed- and open-question survey
VR learning system to support active locomoting viewpoint for astronomy education	(Setozaki et al., 2018)	Knowledge, experience	Assess comprehension of trainee teachers using an astronomical VR learning system	Primary school teachers	Repeated measures	Quantitative: closed-question knowledge test

References

Ampartzaki, M., Tassis, K., Kalogiannakis, M., Pavlidou, V., Christidis, K., Chatzoglidou, S., & Eleftherakis, G. (2024). Assessing the Initial Outcomes of a Blended Learning Course for Teachers Facilitating Astronomy Activities for Young Children †. *Education Sciences*, *14*(6). https://doi.org/10.3390/educsci14060606

Anantasook, S., Kaen University, K., & Chokchai Yuenyong, T. (2015). Thai Students' Understanding about Celestial Motion within Their Social and Cultural Context. In *The International Journal of Science, Mathematics, and Technology Learning* (Vol. 21). www.thelearner.com,

Antoniou, P. E., Mpaka, M., Dratsiou, I., Aggeioplasti, K., Tsitouridou, M., & Bamidis, P. D. (2018). Scoping the window to the universe; Design considerations and expert evaluation of an augmented reality mobile application for astronomy education. In M. Auer & T. Tsiatsos (Eds.), *Advances in Intelligent Systems and Computing* (Vol. 725, pp. 409–420). Springer Verlag. https://doi.org/10.1007/978-3-319-75175-7_41

Aretz, S., Borowski, A., & Schmeling, S. (2017). Development and Evaluation of a Construct Map for the Understanding of the Expansion of the Universe. *Science Education*, *2017*, 1–8. http://edoc.hu-berlin.de/serl

Bartlett, S. (2018). *An Evaluation of Secondary School Students' Learning Experiences with Astronomy-based Physics Outreach Activities*.

Bartlett, S., Fitzgerald, M. T., McKinnon, D. H., Danaia, L., & 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), 1–22. https://doi.org/10.19030/jaese.v5i1.10190

Bartlett, S., Fitzgerald, M. T., Mckinnon, D. H., & Lazendic-Galloway, J. (2018). Astronomy And Science Student Attitudes (ASSA): A Short Review And Validation Of A New Instrument. In *Journal of Astronomy & Earth Sciences Education-June* (Vol. 5, Issue 1).

Benitez-Herrera, S., Spinelli, P. F., Mano, S., & Paula Germano, A. (2019). Pursuing gender equality in Astronomy in basic education: the case of the project "Girls in the Museum of Astronomy and Related Sciences." *EPJ Web of Conferences*, *200*, 02010. https://doi.org/10.1051/epjconf/201920002010

Burrows, A. C., Dipompeo, M. A., Myers, A. D., Hickox, R. C., Borowczak, M., French, D. A., & Schwortz, A. C. (2016). Authentic Science Experiences: Pre-Collegiate Science Educators' Successes and Challenges during Professional Development. *Problems of Education in the 21st Century*, *70*.

Cardinot, A. ;, & Fairfield, J. A. (2019a). *Game-based learning to engage students with physics and astronomy using a board game*.

Cardinot, A. ;, & Fairfield, J. A. (2019b). *Game-based learning to engage students with physics and astronomy using a board game*.

Chen, C. C., & Lin, P. H. (2016). Development and evaluation of a context-aware ubiquitous learning environment for astronomy education. *Interactive Learning Environments*, *24*(3), 644–661. https://doi.org/10.1080/10494820.2014.915417

Chubko, N., Morris, J. E., McKinnon, D. H., Slater, E. V., & Lummis, G. W. (2019). SOLO taxonomy as EFL students' disciplinary literacy evaluation tool in technology-enhanced integrated astronomy course. *Language Testing in Asia*, *9*(1). https://doi.org/10.1186/s40468-019-0095-6

Colantonio, A., Galano, S., Leccia, S., Puddu, E., & Testa, I. (2018). Design and development of a learning progression about stellar structure and evolution. *Physical Review Physics Education Research*, *14*(1). https://doi.org/10.1103/PhysRevPhysEducRes.14.010143

Costa, I. A., Reis, R. C., & Moreira, E. (2023). Science Communication through Astronomy Education: the creation, implementation, and assessment of Porto Planetarium's science education strategy. *Communicating Astronomy to the Public*, *33*, 24–28.

Covitt, B., Friend, D., Windel, C., & Baldwin, J. (2015). A scientific modeling sequence for teaching earth seasons. *Journal of Geoscience Education*, *63*(1), 7–17. https://doi.org/10.5408/14-026.1

Creswell, J., & Plano-Clark, V. (2011). *Designing and Conducting Mixed Methods Research* (2nd ed.). SAGE Publications.

Dankenbring, C., & Capobianco, B. M. (2016). Examining Elementary School Students' Mental Models of Sun-Earth Relationships as a Result of Engaging in Engineering Design. *International Journal of Science and Mathematics Education*, *14*(5), 825–845. https://doi.org/10.1007/s10763-015-9626-5

de Hosson, C., & Décamp, N. (2014). Using Ancient Chinese and Greek Astronomical Data: A Training Sequence in Elementary Astronomy for Pre-Service Primary School Teachers. *Science and Education*, *23*(4), 809–827. https://doi.org/10.1007/s11191-013-9625-2

De Leo-Winkler, M. A., Wilson, G., Green, W., Chute, L., Henderson, E., & Mitchell, T. (2019). The Vibrating Universe: Astronomy for the Deaf. *Journal of Science Education and Technology*, *28*(3), 222–230. https://doi.org/10.1007/s10956-018-9761-1

de Lima, M. S., Vicente Santos, J. C., & F Martin, V. A. (2018). What's That Object? Learning Astronomical Concepts Through The Use Of The Aladin Program And Manipulation Of Astronomical Images. *Journal of Astronomy & Earth Sciences Education-June*, *5*(1).

Dobaria, A., Bailey, J. M., Klavon, T. G., & Lombardi, D. (2022). Students' Scientific Evaluations of Astronomical Origins. *Astronomy Education Journal*. https://doi.org/10.32374/aej.2022.2.1.032ra

Ferrari, E., Teijón, P. H., & Ruiz, C. (2024). Unlocking the cosmos: evaluating the efficacy of augmented reality in secondary education astronomy instruction. *Journal of New Approaches in Educational Research*, *13*(8). https://doi.org/10.1007/S44322-024-00007-8

Görecek Baybars, M., & Çil, M. (2021). Use of portable planetariums in teaching of astronomy: an activity example. *Science Activities*, *58*(2), 43–51. https://doi.org/10.1080/00368121.2021.1906200

Görecek Baybars, M., & Kayabaş, T. B. (2018). Determination of Secondary School 4 th Grade Students' Mental Models of The Astronomer. *International Journal of Evaluation and Research in Education (IJERE)*, *7*(2), 94–99. https://doi.org/10.11591/ijere.v7.i2.11371

Guiotto, L., Fovino, N., Ginolfi, M., Grassi, M., Mascolo, L. Di, & Zanella, A. (2023). *Edukoi: developing an interactive sonification tool for astronomy between entertainment and education*.

Hufnagel, B. (2001). Development of the Astronomy Diagnostic Test. *The Astronomy Education Review*, *1*, 47–51.

Isik-Ercan, Z., Zeynep Inan, H., Nowak, J. A., & Kim, B. (2014). “We put on the glasses and Moon comes closer!” Urban Second Graders Exploring the Earth, the Sun and Moon Through 3D Technologies in a Science and Literacy Unit. *International Journal of Science Education*, *36*(1), 129–156. https://doi.org/10.1080/09500693.2012.739718

Kalkan, H., Kiroğlu, K., Türk, C., Bolat, M., Kalkan, S., & Aslantürk, A. (2014). Basic Astronomy Concepts in the Footsteps of Eratosthenes. *Procedia - Social and Behavioral Sciences*, *116*, 3731–3739. https://doi.org/10.1016/j.sbspro.2014.01.832

Kaplan, A. O. (2014). Research on the Pseudo-Scientific Beliefs of Pre-Service Science Teachers: A sample from astronomy-astrology. *Journal of Baltic Science Education*, *13*(3), 381–393.

Kim, M. S. (2015). Empowering Prospective Teachers to Become Active Sense-Makers: Multimodal Modeling of the Seasons. *Journal of Science Education and Technology*, *24*(5), 610–627. https://doi.org/10.1007/s10956-015-9550-z

Kruse, B., Bass, K. M., & Schultz, G. (2015). The NASA Galileo Educator Network: Using Astronomy to Engage Teachers in Science Practices. In G. Schultz, S. Buxner, L. Shore, & J. Barnes (Eds.), *Celebrating Science: Putting Education Best Practices to Work* (Vol. 500). Astronomical Society of the Pacific.

Kuhn, D., Arvidsson, T. S., Lesperance, R., & Corprew, R. (2017). Can Engaging in Science Practices Promote Deep Understanding of Them? *Science Education*, *101*(2), 232–250. https://doi.org/10.1002/sce.21263

Lee, H., & Feldman, A. (2015). Photographs and Classroom Response Systems in Middle School Astronomy Classes. *Journal of Science Education and Technology*, *24*(4), 496–508. https://doi.org/10.1007/s10956-014-9539-z

Lewis, A. D. (2019). Practice what you teach: How experiencing elementary school science teaching practices helps prepare teacher candidates. *Teaching and Teacher Education*, *86*, 102886. https://doi.org/10.1016/j.tate.2019.102886

Liu, M., Rosenblum, J. A., Horton, L., & Kang, J. (2014). Designing Science Learning with Game-Based Approaches. *Computers in the Schools*, *31*(1–2), 84–102. https://doi.org/10.1080/07380569.2014.879776

McLin, K. M., & Cominsky, L. R. (2014). *NASA Astrophysics Educator Ambassador Program* (J. Manning, M. Hemenway, J. Jensen, & M. Gibbs, Eds.; Vol. 483). http://epo.sonoma.edu/ea/training.php

Mondim, P. (2016). A Week-long Summer Programme in Astronomy for High-school Students. *Communicating Astronomy to the Public*, *21*, 5–9.

Navarro, M. (2014). Evolutionary Maps: A new model for the analysis of conceptual development, with application to the diurnal cycle. *International Journal of Science Education*, *36*(8), 1231–1261. https://doi.org/10.1080/09500693.2013.855343

Plummer, J. D. (2015). *Methods of Engaging Preschool-age Children in Science Practices During Astronomy Activities* (G. Schultz, S. Buxner, L. Shore, & J. Barnes, Eds.; Vol. 500).

Plummer, J. D., Palma, C., Flarend, A., Rubin, K. A., Ong, Y. S., Botzer, B., McDonald, S., & Furman, T. (2015). Development of a Learning Progression for the Formation of the Solar System. *International Journal of Science Education*, *37*(9), 1381–1401. https://doi.org/10.1080/09500693.2015.1036386

Rebull, L. M., French, D. A., Laurence, W., Roberts, T., Fitzgerald, M. T., Gorjian, V., & Squires, G. K. (2018). Major outcomes of an authentic astronomy research experience professional development program: An analysis of 8 years of data from a teacher research program. *Physical Review Physics Education Research*, *14*(2). https://doi.org/10.1103/PhysRevPhysEducRes.14.020102

Rebull, L. M., Roberts, T., Laurence, W., Fitzgerald, M. T., French, D. A., Gorjian, V., & Squires, G. K. (2018). Motivations of educators for participating in an authentic astronomy research experience professional development program. *Physical Review Physics Education Research*, *14*(1). https://doi.org/10.1103/PhysRevPhysEducRes.14.010148

Rodrigues de Andrade, L. (2023). *Astronomy School Education in Chile: Perspectives and Connections from the National Curriculum, Teachers' Content Knowledge and Professional Development Programs*. Pontificia Universidad Catolica de Chile.

Rodrigues, L., Pompea, S. M., Meneses, A., & Montenegro, M. (2024). A Characterization of Astronomy Teacher Professional Development Programs in Chile. *Journal of Science Teacher Education*. https://doi.org/10.1080/1046560X.2024.2375840

Rollinde, E. (2019). Learning Science Through Enacted Astronomy. *International Journal of Science and Mathematics Education*, *17*(2), 237–252. https://doi.org/10.1007/s10763-017-9865-8

Ruddell, N., Danaia, L., & Mckinnon, D. (2016). Indigenous sky stories: Reframing how we introduce primary school students to astronomy - A Type II case study of implementation. *Australian Journal of Indigenous Education*, *45*(2), 170–180. https://doi.org/10.1017/jie.2016.21

Setozaki, N., Tominaga, Y., & Morita, Y. (2018). *VR Learning System to Support Active Locomoting Viewpoint for Astronomy Education* (J. et al. Yang, Ed.).

Slater, S. J. (2014). The Development And Validation Of The Test Of Astronomy STandards (TOAST). *Journal of Astronomy & Earth Sciences Education-December*, *1*(1).

Stork, D. J. (2014). *Contemporary Discipline-based Astronomy Education Research Study of K-12 Teachers' Astronomy Knowledge Using the Test of Astronomy Standards*. University of Wyoming.

Tomita, A. (2019). Daily life astronomy activity and its contribution to children and educators in science education. *EPJ Web of Conferences*, *200*, 02002. https://doi.org/10.1051/epjconf/201920002002

Tsihouridis, C., Mitrakas, N., Batsila, M., & Vavougios, D. (n.d.). A Holistic View of Using Real and Virtual Models in Teaching Astronomy Concepts. In M. Auer, U. Cukierman, E. Vidal, & E. Tovar (Eds.), *Towards a Hybrid, Flexible and Socially Engaged Higher Education* (Vol. 2).

Tsihouridis, C., Mitrakas, N., Batsila, M., & Vavougios, D. (2024). *A Holistic View of Using Real and Virtual Models in Teaching Astronomy Concepts* (pp. 104–115). https://doi.org/10.1007/978-3-031-52667-1_12

Türk, C., & Kalkan, H. (2015a). Developing an Achievement Test for Astronomy Education. *Journal of Studies in Education*, *5*(3), 89. https://doi.org/10.5296/jse.v5i3.8099

Türk, C., & Kalkan, H. (2015b). The Effect of Planetariums on Teaching Specific Astronomy Concepts. *Journal of Science Education and Technology*, *24*(1), 1–15. https://doi.org/10.1007/s10956-014-9516-6

Türk, C., Kalkan, H., İskeleli, N. O., & Kiroglu, K. (2015). Improving Astronomy Achievement and Attitude through Astronomy Summer Project: A Design, Implementation and Assessment. *International Journal of Higher Education*, *5*(1). https://doi.org/10.5430/ijhe.v5n1p47

Wyatt, G., Stevenson, T., & Hamacher, D. W. (2014). Dreamtime Astronomy: Development of a new Indigenous program at Sydney Observatory. *Journal of Astronomical History and Heritage*, *17*(2), 195–204. https://doi.org/10.3724/sp.j.1440-2807.2014.02.06

Zeilik, M., Schau, C., Mattern, N., Hall, S., Teague, K. W., & Bisard, W. (1997). Conceptual astronomy: A novel model for teaching postsecondary science courses. *American Journal of Physics*, *65*(10), 987–996. https://doi.org/10.1119/1.18702

Zhang, J., Sung, Y. T., Hou, H. T., & Chang, K. E. (2014). The development and evaluation of an augmented reality-based armillary sphere for astronomical observation instruction. *Computers and Education*, *73*, 178–188. https://doi.org/10.1016/j.compedu.2014.01.003

Zimmerman, L., Spillane, S., & Sumners, C. (2014). Comparison of Student Learning about Space in Immersive and Computer Environments. *J Rev Astron Educ Outreach*, *1*(1), 5–20. www.eplanetarium.com