

## Chapter 2

# Overview of Planetarium Education Research Methods

The modern digital planetarium, with its inherent ability to show the universe as a real-time three-dimensional model, as opposed to the old-style, earth-centered view of the optical star projector, provides unique opportunities for a fresh look at education research in the planetarium. Now reaching a level of maturity in the second decade of its use, where digital tools are becoming commonplace, the digital planetarium offers a chance to ask old questions with a new perspective that has real datasets of the entire known map of the universe.

Can students identify constellations in the night sky better if they learn in a planetarium or in the classroom? Can students better predict planetary retrograde loops if they have watched planets move using spatially different viewpoints in a digital planetarium? Which astronomy topics are best taught in the planetarium? How much do students enjoy learning in the planetarium? Do classes held in a planetarium improve attention and retention? Do visits to the planetarium increase the number of students entering the science and engineering pipeline? How do students make sense of the celestial sphere? These, and countless other questions, dominate the realm of planetarium education research.

Sometimes asking these questions results in refereed journal publications or in doctoral dissertations that add to the academic literature base stored in university libraries. Other times, asking these questions results in evidence-based talks given at professional planetarium society conferences that consequently urge commercial planetarium program developers to dramatically change the way that information is presented to audiences. These developers come up with new user friendly tools, which make learning in the planetarium more durable and effective. And, sometimes, like in much of scientific inquiry, study results end up creating more questions than actually providing solid actionable answers. This chapter provides a brief introduction to the major methods and underlying foundations adopted by the broad community of scholarly planetarium education researchers in their academic pursuits.

Planetarium education research typically falls under the broader scientific domain of discipline-based education research. Discipline-based education research, as defined by Slater (2015a, b) and colleagues, is the scientific enterprise of making systematic observations and conducting ethical experiments of learners and their environments in order to develop predictive and explanatory models of teachers, teaching, learners, and learning, in the service of enhancing learning. Such a definition is consistent with other authors, including those representing the National Research Council and the National Academy of Sciences (Singer et al. 2012). Although some traditional scientists occasionally mock the science education research enterprise as being subjective and lacking the objective rigor of conventional bench-based laboratory science as a whole, a systematic and data-driving scientific approach to better understanding the process of teaching and learning is a respected and formally supported part of the larger scientific community. The largest professional scientific societies have all issued position statements emphasizing the value of carefully conducted scientific research in the domains of teaching and learning. Figure 2.1 lists some of these position statements.

There are a wide variety of domains in which a planetarium education researcher might apply their scientific inclinations to better understand the role planetariums play in the educational landscape. One could study how people learn from the planetarium, for certain. And, one might also choose to study how people feel after a planetarium learning experience. Additionally, one might choose to compare the type of planetariums in formal school-building environments to those in museum and science center environments. Alternatively, one might choose to compare the nature and specific scientific content in the programming of permanent brick-and-mortar planetarium buildings versus the inflatable portable planetariums. One might even choose to study the range and domain of the academic qualifications of individuals working in the planetarium field. What about studies on learning via computer desktop planetarium simulation programs on a flat screen monitor? Given the seemingly endless range of possibilities, it seems necessary, albeit unappetizing, to constrain planetarium education research to some degree in order to begin making sense of it.

One of the most straightforward approaches to constraining the range of planetarium education research efforts might be to separate the vast number of study possibilities such that studies fit into one of two categories: descriptive and experimental. On one hand, descriptive studies provide an objective, third-party narrative of what happened during an educational event, the sequence of how it happened, and why the event took place. Such studies can provide long descriptions or conduct a comparative analysis of different educational situations. An example would be to describe the number and ages of people who visit a planetarium. Another example would be to survey planetarium directors across the country and tabulate which show schedule times attract the largest number of attendees.

On the other hand, experimental education studies attempt to measure a change and attribute that change to an educational intervention. There are at least two traditional ways to conceptualize a basic experimental study. One way is to take a group of study-participants and measure something about them—their knowledge

---

American Astronomical Society Position Statement on Research in Astronomy Education  
Adopted 2 June 2002

<https://aas.org/governance/council-resolutions#edresearch>

In recent years, astronomy education research has begun to emerge as a research area within some astronomy and physics/astronomy departments. This type of research is pursued at several North American universities, it has attracted funding from major governmental agencies, it is both objective and experimental, it is developing publication and dissemination mechanisms, and researchers trained in this area are being recruited by North American colleges and universities.

Astronomy education research can and should be subject to the same criteria for evaluation (papers published, grants, etc.) as research in other fields of astronomy. The findings of astronomy education research and the scholarship of teaching, when properly implemented and supported, will improve pedagogical techniques and the evaluation of both teaching and student teaching.

The AAS applauds and supports the acceptance and utilization by astronomy departments of research in astronomy education. The successful adaptation of astronomy education research to improving teaching and learning in astronomy departments requires close contact between astronomy education researchers, education researchers in other disciplines, and teachers who are primarily research scientists. The AAS recognizes that the success and utility of astronomy education research is greatly enhanced when it is centered in an astronomy or physics/astronomy department.

---

American Physical Society Policy Statement on Research in Physics Education  
Adopted May 21, 1999

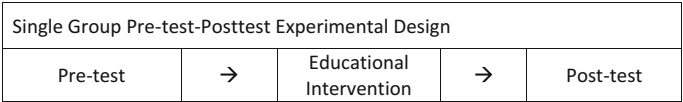
[https://www.aps.org/policy/statements/99\\_2.cfm](https://www.aps.org/policy/statements/99_2.cfm)

In recent years, physics education research has emerged as a topic of research within physics departments. This type of research is pursued in physics departments at several leading graduate and research institutions, it has attracted funding from major governmental agencies, it is both objective and experimental, it is developing and has developed publication and dissemination mechanisms, and Ph.D. students trained in the area are recruited to establish new programs. Physics education research can and should be subject to the same criteria for evaluation (papers published, grants, etc.) as research in other fields of physics. The outcome of this research will improve the methodology of teaching and teaching evaluation.

The APS applauds and supports the acceptance in physics departments of research in physics education. Much of the work done in this field is very specific to the teaching of physics and deals with the unique needs and demands of particular physics courses and the appropriate use of technology in those courses. The successful adaptation of physics education research to improve the state of teaching in any physics department requires close contact between the physics education researchers and the more traditional researchers who are also teachers. The APS recognizes that the success and usefulness of physics education research is greatly enhanced by its presence in the physics department.

---

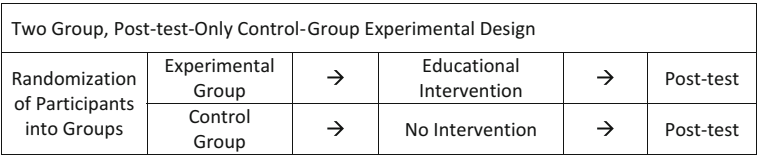
**Fig. 2.1** Examples of professional scientific society statements on education research



**Fig. 2.2** Schematic of single group pre-test-posttest experimental design

or their attitudes for example—before and after an educational intervention, and determine how much change has occurred (Fig. 2.2). Another way is to take a sample of study subjects and randomly assign them to different treatments and measure how groups differ as a result of their various assigned treatments (Fig. 2.3). It is important to emphasize that study subjects must be randomly assigned to various treatment groups to be a true experiment. The reality of working with school children, for instance, is that it is incredibly difficult to randomly assign students to different treatments. Instead, what usually happens for pragmatic reasons of scheduling and supervision is that students are not randomly assigned to different treatments. This results in a study being classified not as experimental, but as quasi-experimental. Being classified as a quasi-experimental research study is not necessarily a pejorative label; rather, the label is an accurate description of how a study was done as a result of how participants are selected and assigned treatments.

Unceremoniously dividing the breadth of planetarium education research work into just two broad categories is probably insufficient to capture the breadth and depth of the field. An early scholarly effort to constrain what constitutes the domain of education research was advocated by Novak (1963). Novak (1963) felt strongly that science education research is best done by people who are inclined toward, and perhaps formally trained in, scientific observation and experimentation. Underlying his thinking was that educators of the day were naturally focused on nurturing youth as growing and evolving humans and would be largely unable to make rationally objective and evidence-based conclusions about comparative science education interventions. More than who should be doing science education research, he also argued passionately that the role of science education research should be entirely focused on advancing a priori explanatory theory through experimentation. In other words, the goal of science education research should be to develop sufficiently complete educational theories through repeated experimentation so that one could successfully predict the educational outcome of a proposed educational intervention without actually doing the experiment.



**Fig. 2.3** Schematic of two group, post-test-only control-group design

Novak (1963) divided the landscape of science education research into four domains: baseline studies, correlation studies, experimental studies and curriculum studies. *Baseline studies* are those census surveys of the field that establish a snapshot of the current status and characteristics of planetarium education research that are normative descriptions of the field: How many planetariums are there? Who attends planetarium shows and what time of day? What is the portfolio of academic qualifications of planetarium lecturers? These studies are important for individuals who want to watch the evolution of planetarium education and understand the leverage points of engaging potential populations of new attendees.

A second category of studies Novak (1963) identified are *correlation studies*. Correlation studies are those in which two or more variables are compared to determine how often they occur together. An example question would be: do students have more achievement gains when learning in larger sized domes, as compared to smaller domes? Correlation studies are useful, but must be interpreted with a degree of caution. An often repeated mantra in planetarium education research is that correlation does not equal causation. In the example cited, larger domed facilities, such as those in museums and science centers, might be more often visited by students representing a distinct socio-economic group as compared to smaller domes. This might account for the differences in achievement gains seen in pre-testing and post-testing visiting students. In education research, a participant's socio-economic status can have a major influence on achievement and attitude changes as a result of an educational intervention. Such a scenario requires researchers to carefully describe the characteristics and demographics of any and all study participants.

The most powerful research studies that Novak (1963) identifies are scientific *experimental studies* (viz., Figs. 2.2 and 2.3). These are education research studies where all, or most, variables are identified and only one independent variable is carefully manipulated at a time. All other variables are held constant, or randomized, so that a responding, dependent variable can be measured. In this case, the responding, dependent variable can be reasonably attributed to changes in the manipulated, independent variable. For example, if one group of students sees a planetarium show that includes significant humor, and another group of students sees the same show without humor included, one can reasonably attribute any observed differences in learning gains to the inclusion or exclusion of included humor if the initial equivalency of the two groups can be firmly established (e.g., Fisher 1997). In this illustrative example, as in many others, the demographics of any participants in the study need to be carefully described. Changes in achievement due to the inclusion or exclusion of humor in a planetarium program may be the result of differences in cultural background, so this needs to be accounted for. Furthermore, each population of students may provide varying results because of differences in cultural backgrounds.

A final category of studies that does not fit into the three previously identified by Novak (1963) is what he terms *curriculum studies*. This category includes studies on the astronomical topics we teach in each facility and at each grade level. For example, in the US, one might be interested in describing how the *Next Generation*

*Science Standards* are being manifested in planetarium education (Schleigh et al. 2015). One of the problems with conducting this sort of academic work is that authors identify what is being taught in education systems by popularity due to high frequency, rather than what is established as core ideas of a discipline. Planetarium education systems in the US have struggled with this over recent decades, in particular, because of generous funding from NASA. These funds to education and outreach programs enhance the public's support of particular NASA science and specific NASA missions, rather than astronomy and space science as a coherent whole. Novak (1963) argued that curriculum studies should only be taken on by individuals with an exceptionally strong background and deep understanding of the philosophy of science and the historical development of scientific disciplines. More critically, Novak (1963) argued that curriculum studies are most often unworthy of being reported in the archival scholarly literature as they are generally focused at the local level where the study was conducted. They also usually have little or no value to the larger science education community, unless some promising new research method is utilized to conduct the curriculum study.

Ball and Forzani (2007) advocate a vision of education research that offers guidance in terms of the type of studies the planetarium education research community could most reliably include as their domain of education research. In response to the value of education research studies too often being questioned by skeptics, well-respected, education thought leaders Ball and Forzani (2007) argue that the education researcher—and in the current case, the planetarium education researcher—would be best served by focusing their scholarly studies on research within education, instead of studying phenomena that is only related to education. In this sense, examples of study on topics related to education would be counting and classifying the number of planetariums by dome size or the reasons for people returning to see each and every planetarium program a facility presents. These are not key aspects to understanding the underlying mechanisms of planetarium education.

Ball and Forzani (2007) instead argue that the planetarium education research that really matters is that which focuses on what they call *instructional dynamics* and *educational interactions*. In this sense, they encourage the community's scholarly researchers to concentrate on how planetarium programming can uniquely interpret and present astronomy to students. They also support research on the specific mechanisms that drive students' responses to that stimulus either in terms of achievement or attitude. An example consistent with this perspective in planetarium education research would be which specific learning interventions occurring in the planetarium cause changes in students' understanding—be they hands-on with manipulates or minds-on with cognitively challenging mental engagement tasks. These are characterized by studying interactions between the planetarium stimulus and students' responses. Some examples that are inconsistent with this modern idea of worthy research are studies that comprehensively investigate students' attitudes or beliefs about astronomy or career aspirations, but provide no insight on the educational transactions at planetariums that influence students' thinking. In other words, planetarium education research that identifies changes in

students, but is unable to robustly establish causal mechanisms, falls far short of Ball and Forzani's (2007) vision.

## Ethical Guidelines for Education Research

What was judged appropriate planetarium education research fifty years ago, and what will be considered the most useful planetarium education research going forward, is evolving substantially. One of the monumental changes occurring in planetarium education research over the last two decades is an enthusiastic national focus on researchers being highly ethical in human studies. In the 1960s and 1970s, seminal planetarium education research studies were done by surveying students upon entering and exiting the planetarium. The question about whether or not these early era study participants were willing subjects was not really a concern for most planetarium education researchers. In those days, students were relatively compliant and when an authority figure in education, such as a teacher or a planetarium lecturer, asked students to voluntarily complete a survey without compensation, students most often acquiesced to the request without complaint from either themselves or their parents. This situation was not unique to planetariums, but was characteristic of individuals in most educational systems. In essence, when students would enter the planetarium, or classroom, it seemed perfectly appropriate to give them a survey about their understanding of, or attitudes toward, astronomy and expect them to complete the survey.

A perspective that data could be collected from students without their well-informed understanding and consent has changed dramatically in the last two decades. Today, planetarium education researchers are bound by ethical guidelines that require human subjects to be fully informed of any study they are participating in, no matter how insignificant it might appear to the researcher (Slater et al. 2015a). Normally, this is done by verbally describing the study and its potential benefits and risks. No study has zero risks. A participant could unexpectedly become anxious or distraught while completing a survey if they do not know the answers. More importantly, no study has absolutely zero costs. If you require a participant to complete a five-minute survey, that is five-minutes of their personal disposable time that they could have used in another way of their own choosing, perhaps responding to a critical email instead of completing a survey. Also, asking participants to complete a survey in the planetarium could take away time that could have been spent actually learning something from the planetarium. In other words, there are risks and costs to being a research participant that novice planetarium education researchers might not realize. Most often, voluntary participation is solicited and consented to when researchers provide both an orally spoken description and a written consent form to be signed by the participant. An illustrative example is given in Fig. 2.4. Parents or guardians may also need to sign a consent form if study participants are not of legal age (18 in most states).

My name is \_\_\_\_\_ and I want to tell you about a research study I am doing. A research study is usually done to find a better way to understand how things work. In this study, I want to find out more about the best way for students to learn \_\_\_\_\_.

The research study you will participate in will take place \_\_\_\_\_. During the research study you will \_\_\_\_\_. The lesson and assessments should take no longer than \_\_\_\_\_.

There is minimal risk involved in the participation of this research study. There is a chance you may feel anxiety before, during, or after the pre- and post-assessments, or you may become embarrassed during the lesson if you do not know the answer to questions I ask. But these risks are minimal and any embarrassment, anxiety, or discomfort that this research may cause is no greater than that of any other risk that you may encounter in a classroom, during a lesson, or while taking a test. If you do feel anxiety or discomfort at any time just let me or your teacher know and we will allow you to take a small break or exit the room.

There are benefits for you to participate in this research study. The main benefit is an increased knowledge on \_\_\_\_\_, which is part of your curriculum. The knowledge gained during this research study will also be helpful for your in-class and standardized statewide tests. This study also has the potential to help teachers and curriculum developers to better teach future students. All of your data will be kept confidential. Each of you will be identified using a numerical code, which will be given to you during the initial meeting. You will use this numerical code on all assessments during the study. No identifying characteristics will be used in the presentation of the data. If you are referred to in any class assignment or publication, only your pseudonym will be used.

I will not discuss your name with anyone other than your teacher or my colleagues. All data, including signed consent and assent forms, master lists with your name and corresponding numerical code, pre- and post-test scores, and spatial ability scores will be stored in a digital format on a password protected hard drive that only myself and my colleagues can access. All hard copies will be stored in a locked filing cabinet in my home office that only I will have a key to. The research summary, all tests, signed consent forms, and signed assent forms will be stored for three years after the research is completed, until November 30, 2019. Your data may be used for future class projects I am assigned or future publications, but again, your identification will be confidential and no names will be provided. Your participation is voluntary and your refusal to participate will not involve penalty or loss of benefits to which you are otherwise entitled, and you may discontinue participation at any time without penalty or loss of benefits to which you are otherwise entitled.

Also, you understand that your refusal to participate or your withdrawal at any point will not affect your course grade or class standing.

If you choose not to participate in this study at any time please let your teacher or myself know. You will still be expected to participate in the lessons as part of your science curriculum but no data from your surveys will be collected or recorded. I am asking to use this data for research purposes.

If you have any questions or concerns about participating in this research study please contact myself \_\_\_\_\_ or my supervising colleagues at \_\_\_\_\_.

\_\_\_\_\_  
PRINT NAME

\_\_\_\_\_  
SIGNATURE

**Fig. 2.4** Illustrative example of a human subjects consent signature form for minors

There are times and situations, however, when a particular study design benefits from a scenario in which participants can be deceived briefly about the nature of a study by the researcher, but such situations are nuanced, delicate, and rarely are part of the planetarium education research world. An oversimplified example of a



rapidly conceived research study on the role of narration during a planetarium program might be where English-only speaking planetarium visitors are deliberately misled and misinformed that the only narrated soundtrack available is in Spanish, and that they will have to learn astronomy as best they can in this less than advantageous situation. The bottom line is that participants can only consent to be included in a research project when they are fully informed about the potential benefits and risks.

The types of experimental studies briefly mentioned earlier and illustrated in Figs. 2.2 and 2.3, can vary considerably in terms of specific configurations. However, nearly all research designs in planetarium education research surround the same theme: what was the impact of visiting the planetarium? The five most common study designs to uncover what happens as a result of learning in the planetarium are illustrated in Fig. 2.5.

The most basic study design is the snapshot case study. In this study design, planetarium attendees are surveyed after they have visited the planetarium. The advantage of this particular design is that it is relatively easy to administer—often in the form of an exit survey. The disadvantage is that the researcher does not know what the participants’ responses were before they attended the planetarium and, as such, cannot robustly know that the planetarium experience itself had any impact at all. Instead, a stronger research design is the single-group, pre-test—post-test design. In this case, a researcher is able to establish what the incoming responses

**Fig. 2.5** Illustrations of common study designs

R = Random Placement of Participants into Groups;  
EI = Educational Intervention;  
O/M = Observation/Measurement

*ONE:* Snap-Shot Case Study

— EI O/M

*TWO:* Single-Group Pre-test-Post-test Design

O/M EI O/M

*THREE:* Pre-test-Post-test Control Group Design

R O/M EI O/M  
R O/M — O/M

*FOUR:* Post-test-Only Control Group Design

R — EI O/M  
R — — O/M

*FIVE:* Solomon Four-Group Design

R O/M EI O/M  
R O/M — O/M  
R — EI O/M  
R — — O/M

are from a pre-test and any differences in responses on a post-test can be roughly attributed to the planetarium learning experience itself. The disadvantage to this design is that one cannot know the impact of taking the pre-test. For example, it is possible that taking the pre-test unsuspectingly telegraphs or cues participants to pay more attention to some parts of the planetarium learning experience than they otherwise would have. In this case, the experience of taking the pre-test itself becomes an uncontrolled, confounding variable.

## Control Group Design

To mitigate for the potential impact of taking a pre-test on the study participants' post-test responses, most planetarium education researchers adopt the pre-test—post-test, control group design. This study design is characterized by randomly dividing the study population into two groups, where one receives the educational intervention, and the other does not. In addition to mitigating for any unknown impacts of taking the pre-test on study participants, this design also allows the researcher to establish an initial equivalence among the two groups. If the two groups are not initially equivalent, results could be unknowingly influenced by one group that is already educated in the planetarium or previously had their attitudes positively enhanced. This research design has the further advantage of being able to more confidently attribute differences in post-test scores to a specific educational intervention in the planetarium compared to the two previously mentioned designs.

A particular characteristic of the pre-test—post-test control group design is worth noting, and it is related to the term *control group*. Planetarium education researchers often describe two comparison groups of study participants as either being members of the *experimental group* or the *control group*. Most often, the study participants that were described as the experimental group was the group that learned in the planetarium and the group of study participants that learned in the traditional classroom is called the control group. In modern writing, planetarium education researchers should be most appropriately referring to the two involved comparison groups as *treatment-1* and *treatment-2* groups. This is because in a traditional, scientific experimental research study, one of the groups should receive a treatment—in this case an educational intervention—and the other group should receive no educational intervention treatment at all, if it is actually a control group in the strictest sense. Very rarely in planetarium education research do we read about studies with a non-treatment control group that received no educational intervention whatsoever, even though this is how they are sometimes described. There is an ethical reason for this as well. Most planetarium educators do not want to academically disadvantage half of the students by not giving them any instruction in astronomy. As a result, researchers either give students an alternative learning experience or, what is now more commonly done, provide the non-planetarium learning group a chance to go to the planetarium after the experimental study has concluded.

Depending on the specific nature of the study, even the pre-test—post-test control group design does not fully account for the impact of participants taking a pre-test. The most well-known problem related to pre-tests has to do with studies involving tests of spatial reasoning. It has been observed that scores improve each time students take a spatial reasoning test, even when there is no educational intervention. In other words, simply taking a spatial reasoning test improves your score the next time you take a spatial reasoning test (Heyer et al. 2013). Somehow, the act of taking the test increases your ability to do spatial reasoning. In response, some researchers argue the advantages of using a post-test-only control group design over the pre-test—post-test control group design, depending on the precise nature of the research question.

Perhaps the strongest research design available to planetarium education researchers is known as the Solomon four-group design. Illustrated in Fig. 2.5, the Solomon four-group design contains two extra control groups, which allows the researcher to explicitly test the extent to which the pretest itself has an effect on study-participants' post-test responses. Admittedly, this powerful research design is much more complex and requires considerably more effort on the part of the planetarium education researcher to conduct. Campbell and Stanley (1963) convincingly argue that when the Solomon Four-Group Design is impractical, the post-test-only control group design is a stronger design than the pre-test—post-test control group design. It is more powerful to argue that two groups were initially equivalent than to argue that a pre-test had negligible impact on the participants prior to an educational intervention.

## Validity and Reliability

In any study, understanding the nuanced strength and weaknesses of the measurement instruments themselves is usually the most challenging. In all cases, researchers must make a compelling argument that the instruments used are both reliable and valid. The concepts of reliability and validity can be likened to synonyms for precision and accuracy, respectively. Reliability refers to “being consistent in the sense that a subject will give the same response when asked again”. That is to say, if one was to step on a scale multiple times and it gave results such as 135, 134, 134, and 135, one could say that that instrument was reliable. However, if one stepped on the scale and the results were much more varied (e.g., 128, 142, 136, and 112), then this would show that the instrument was unreliable and could not be counted on to give precise, consistent answers. Validity, on the other hand, has more to do with measuring “what is intended to [be] measured and accurately reflecting the concept”. Extending the same example, if one knows that the weight should actually be 115 lbs but the scale is still reading 135, 134, 134, and 135 there must be some sort of systematic error.

In order to test for reliability of an assessment tool, there are a few different methods such as test-retest, equivalent-forms, internal consistency, and interscorer,

all of which calculate a correlation coefficient to determine its reliability. The test-retest reliability method is computed by administering a test over time; equivalent-forms reliability uses multiple forms of a test created to measure the same construct; interscorer reliability is determined by having multiple scorers agree on the results given; and internal reliability refers to “how consistently the times on a test measure a single construct or concept”. Most tests are supposed to measure a one-dimensional construct, and tests that do this have high internal reliability. However, other tests, such as the *Test Of Astronomy Standards* (TOAST), are multidimensional so in order to determine its internal reliability, each construct must be measured separately (Slater 2015). To calculate the internal reliability, researchers use the Chronbach alpha coefficient, most commonly called the “coefficient alpha” (Chronbach 1951). The coefficient alpha is calculated using the number of items on a test and the average correlation coefficient. A coefficient alpha higher than 0.7 is typically considered internally reliable and anything less than that could be due to test fatigue or some other underlying cause.

Validity can be a bit more difficult for which to argue. However, validity is just as important as reliability. As Chronbach says in his 1991 article, “a test may be excellent in other respects, but if it is wrongly interpreted it is worthless in that time and place”. Ensuring validity is done by gathering evidence that supports the inferences and findings made by the researcher. Validation is one of the most important reasons a research study should be seated firmly within the context of the field and should “stand on the shoulders of giants”. There are four main types of validation: construct, face, criterion, and content validity. Most validation starts out with face validity, by the researcher simply saying to his or herself, as well as asking an outsider’s expert opinion, “it looks like it’ll work”. While it is important for the researcher to reflect on how valid their assessment tool seems, it is highly subjective and requires other forms of validation. Construct validity is determined by establishing the test measures the construct was designed to measure. For example, if an exam was trying to evaluate students’ understanding of astronomy concepts, the developer should ensure that the reading level is well below the students’ grade level to make sure reading ability does not affect the results. Criterion-related validity uses tests that have been previously validated for measuring a similar construct, to see if the results correlate with those outcomes. Lastly, content validity determines if the content used on the test is appropriated considering the subject area that is being assessed. For example, if a researcher wants to know how well students understand lunar phases, asking a question about celestial motion may be irrelevant. This type of validation should be done by asking experts in the field.

More specifically, a researcher may wish to determine the difficulty of certain items and evaluate which questions students “who get it” get correct versus questions students “who don’t get it” get correct. These can be measured by calculating item difficulty and item discrimination. Item difficulty is a value given from 0 to 1 and equates to the percentage of students who answered an item correctly. If an item has a difficulty of 0.30 and a high discrimination score, then the students who did well on the overall exam are likely to make up the 30% of students who

answered the question correctly. If the discrimination is low or negative for this item, it means that students who did not do well on the exam probably guessed on this item. Item discrimination can be determined by calculating the biserial-R. It is a correlation statistic from  $-0.1$  to  $+1.0$ . If the score is near 0 it shows that answering that question “has nearly no relationship to overall matter of the material”. Negative values indicate that students who did well on the overall exam did poorly on that particular item. These types of questions should be analyzed for flawed construction (Slater and Adams 2003).

When analyzing the quantitative and mixed methods studies submitted to the iSTAR database, we prefer articles that have clear and specific methods, as well as results sections explaining the procedures used to calculate whether results are reliable, validated, and significant. To ensure reliability, especially of a self-made assessment instrument for a specific study, we prefer to see a coefficient alpha of at least 0.7 or higher to ensure internal reliability. An additional reliability test would also help, such as the test-retest or interscorer methods. As for validation of a test instrument, numerous methods of validation are important. Construct validity and content validity ensures the researcher is measuring what they intend to measure and use appropriate content-related items. Criterion-related validity can be a more objective voice to the validation process, in order to determine the new assessment instrument’s correlation with a previously deemed valid and reliable instrument. Face validity ensures that the overall structure and construction is sound. As reliability can be shown statistically, validation should be discussed in length in a study and it should be clear that the researchers went to exhaustive and extensive length to prove the validity of their instrument. Due to the complex nature of designing an instrument, it is typically best if researchers use a previously constructed instrument that has already been checked for reliability and validity. Hence the reason instruments such as the TOAST are so important for each field. Universally used instruments in a field also benefit the literature by providing a way for researchers to compare the outcomes of their studies. These instruments can also provide insight into the effectiveness a variety of dimensions that novices should be learning about in introductory astronomy courses based on different instructional methods, curriculum, and activities.

One example showing the importance of validity is a classic planetarium education research study done by Akey (1973). Akey (1973) conducted a pre-test—post-test study and found statistical significance in students’ learning gains from using the planetarium on 56 astronomy concepts. However, reflecting on his assessment instrument today shows a lack in face validity. First, the test’s reading level for the second grade students was too difficult. Also, 56 concepts would now be considered far too many astronomy concepts to teach such a young group of students in that short amount of time. Lastly, students were able to answer a large percentage of these answers correctly on the pre-test which is highly unlikely given what other research on planetarium education stated about the students’ understanding of astronomy at that age level. Due to these validation issues, researchers now know that using the results from a study like this would be unscholarly and futile. The ability to reflect on the validation and reliability of previous studies is

critical to understanding the quality of research that can truly help fuel the future of astronomy education research.

Even when the validity and reliability of an instrument has been thoroughly discussed, there are situations when these can be threatened by confounding issues beyond the survey instrument itself. Examples include when major newsworthy events occur between observations, when exhaustion (or maturation) occurs during or between observations, when observers or scorers standards drift from grading survey answers early in the study to late in the study, unknown and unintentional selection bias of selected participants perhaps from a particular socio-economic group, and last, but not least, interactive effects due to the mere presence of survey instrument or an observer in the room taking field-notes.

### Bloom’s Taxonomy

Almost all survey instruments, regardless of their form, attempt to quantify not just the amount of learning that occurs or the extent of attitude change, but the actual depth and quality of that learning. Most planetarium education researchers lean heavily on the longstanding theoretical work done decades earlier by Bloom (Anderson et al. 2001; Bloom 1956; Krathwohl et al. 1964). Bloom’s (1956) initial efforts were based on the need to classify knowledge survey questions in terms of how many cognitive resources were required to answer them correctly. He started by classifying questions from a ranging of relatively easy recall questions and, with increasing cognitive complexity, to highly complex questions requiring the synthesis of numerous ideas in order to evaluate novel situations. Eventually, he concluded that assessment questions—and their corresponding learning objectives—could be categorized consistently into six cognitive levels. Known broadly as the *cognitive domain*, Bloom’s six-level taxonomy ranges from the lowest cognitive levels to the highest cognitive levels in the sequence: (1) Knowledge; (2) Comprehension; (3) Application; (4) Analysis; (5) Synthesis; and finally, (6) Evaluation. In Fig. 2.6, each label refers to the cognitive tasks required to complete a task successfully. Note

Bloom’s Taxonomy of Cognitive Domain Learning Objectives		
Levels	Names	Characteristics
Level One	Knowledge and memory	Identify, list, repeat, recognize
Level Two	Comprehension and understanding	Locate, represent, categorize, explain
Level Three	Application	Relate, calculate, organize, use
Level Four	Analysis	Compare, contrast, experiment, discriminate, structure
Level Five	Synthesis	Compose, create, generalize, relate, construct
Level Six	Evaluation	Judge, estimate, criticize

**Fig. 2.6** Bloom’s taxonomy of cognitive domain learning objectives

that each label has evolved somewhat over the decades. Successfully identifying the brightest star in the night sky as Sirius would be classified as a knowledge-level task, and evaluating a prediction of Earth’s observed retrograde loop duration as seen from the surface of Mercury would be a synthesis-level task.

The precise language that planetarium education researchers use to describe Bloom’s Taxonomy varies somewhat from one person to the next. While one researcher might refer to the knowledge level as the lowest level and the evaluation level as the highest level, another researcher might use words vaguely suggesting that the levels are inverted by ascribing the knowledge level as being the shallowest level and the evaluation level as being the deepest level, rather than the highest level. Be that as it may, Bloom’s Taxonomy has been a highly valuable part of planetarium education researchers being able to describe the specific depth of understanding that they are testing and relating the impact that planetariums have on learning.

Many planetarium education researchers are also often interested in the evolving attitudes, values, interests, feelings, and motivations of people engaged in learning in the planetarium. Along with Krathwohl et al. (1964) and colleagues, Bloom also advanced a taxonomy describing the *affective domain*. A learner that has achieved affective learning at the lowest level of the affective domain is an individual who is clearly open to new experiences and willing to devote emotional energy to a person, environment, or event. In contrast, an individual who has achieved the highest level—or deepest level, according to some—is one who has internalized their feelings to the point that they will obviously behave in a predictable way, responding to their well-developed worldview. Bloom’s definitions of the affective domain are not as widely adopted in the planetarium education research community as those descriptions of the cognitive domain. Nonetheless, Bloom’s taxonomy of the affective domain is much more precise for describing attitudes rather than just the participants’ high or low attitudes and levels of interest in a topic (Fig. 2.7).

Understanding the process of learning in a planetarium in terms of their specific depth of comprehension—cognitive domain—and characterizing their wealth of attitudes, values, and interests—*affective domain*—is indeed complex. Adding to this complexity is the environmental context where learning in the planetarium

Bloom’s Taxonomy of Affective Domain Learning Objectives		
Levels	Names	Characteristics
Level One	Receive	Open to experience, attend, acknowledge
Level Two	Respond	React, respond, clarify, reference
Level Three	Value	Judge relevance, debate, confront, justify
Level Four	Organize	Personalize, defend, prioritize, contrast
Level Five	Internalize	Act, display, practice, adopt a behavior

Fig. 2.7 Bloom’s taxonomy of affective domain learning objectives

occurs. In the broadest terms, planetarium learning experiences can occur in formal, school-based learning environments or in informal, out-of-school learning environments. In abbreviated form without any regard to value or hierarchy, planetarium education researchers generally use a short-hand description of learning environments as either being formal or informal (Falk and Dierking 2012).

## Learning Environments

Both formal and informal learning environments are deeply critical to the education landscape at large (Feder et al. 2009). Besides the physical location being different between formal and informal learning environments, they also differ to greater and lesser extents in physical size, setting, populations, targeted learning objectives, and ways of assessing their effectiveness. Conducting planetarium education research in these domains can appear vastly different in the methods used for collecting data and how one interprets collected data.

A formal learning environment is typically set within the classroom or at an academic institution, where students are evaluated on a regular basis on their learning progressions by receiving feedback through grades or tests. Traditionally, this type of environment is controlled and structured with instructors “feeding” information to the students through instructional strategies such as lectures, group/individual activities, laboratory demonstrations, etc. During each class, students are expected to participate in scheduled lessons laid out in the syllabus. Additionally, formal environments are set over long, repetitive periods of learning designed to develop a student’s understanding through a pre-designed learning progression embedded in the curriculum.

Informal learning environments, in contrast, have a wider range of settings in which they occur, such as museums, after-school programs, nature centers, or social settings (Falk et al. 2007). Feedback received in these types of environments is typically either immediate or does not occur at all. Usually, there is no opportunity for the individual to practice retrieving the information unless they choose to in another setting. Conventionally, these types of environments are unstructured and allow the learner to explore along a path of their choice. Other than being semi-guided by the constraining floor plan of an exhibit hall at a museum, it in no way ensures how much time a learner will spend at each exhibit nor how deeply they will engage with the content.

## Population

Let’s first consider the population differences. An informal learning environment consists of a wide range of age groups, from pre-kindergarten to those well into retirement. This type of audience creates a challenge for those responsible for



engaging such a diverse collection of learners. Each age range has a different set of generational and life-stage interests. For example, a young group of learners could require a more technologically advanced, hands-on type of activity, whereas those in retirement may be more comfortable and able to read a long script of information in silence. In addition, informal environments are charged with the challenge of potentially extremely diverse cultural populations. As important as it is to use a culturally relevant mode of instructional strategies in the formal classroom, it is exceedingly more difficult to define those margins of characteristics in an informal environment where the same features can change daily.

Unlike informal learning environments, a formal setting is expected to have a narrow range within the students' age. The advantage of formal settings is that they can create a curriculum that is specifically designed to target their audience's cognitive abilities. By understanding what is most relevant to students at their age-level, instructors are able to design activities and use instructional strategies that are optimal for engaging their learners. Generally, formal learning occurs between the ages of 5 and 18 years and, for those who choose to continue in higher education, up to approximately 30 years of age. Formal education forces the learner to engage with particular concepts they may not have otherwise. Teachers in formal environments can guide students to the correct answer and provide critical feedback when the learner's current model about a concept is incorrect.

## **Objectives and Assessment**

Formal and informal learning environments also differ in terms of the targeted learning objectives and the evidence they collect for assessment of outcomes. The National Research Council study committee on informal education describes a variety of outcomes that should be central to informal environments, such as promoting lifelong, life-wide, and life-deep learning (Feder et al. 2009). Specifically, this study committee created a list of six strands of outcomes from which informal outcomes should operate their framework. The six strands include developing an interest in science, understanding science knowledge, engaging in scientific reasoning, reflecting on science, engaging in scientific practices, and identifying with the scientific enterprise. In order for researchers to assess whether or not these outcomes are being achieved, the scope of assessment tools can be limited due to the unstructured nature of informal settings. Generally, self-reporting assessments tend to be the most common, if not one of the only ways that researchers are able to measure these objectives. Modes of self-reports include, but are not limited to, questionnaires, structured interviews, surveys, focus groups, and talk-aloud protocols.

## Informal Setting

In particular, one of the most difficult strands of outcomes to measure in informal environments is the level of understanding scientific knowledge. The reason for this difficulty is due in part to the individualistic nature of this strand and the fact that learners come to the setting with prior knowledge, understandings, and experiences. Evaluating these gains requires thorough pre- and post-tests or interviews, which can be challenging to obtain as well as time consuming.

Due to the challenges that informal learning researchers face, affective variables, such as motivation and interest, are often the primary focus of informal learning research. Researchers such as George Reed, from the planetarium education research community, have found overwhelming evidence in their respective field that the use of the planetarium versus the use of classroom to learn astronomy holds no inherently superior advantages. Influenced heavily by this finding, Reed (1973), and Ridky (1974) advanced the notion that perhaps the greatest value that the planetarium holds is within the affective domain.

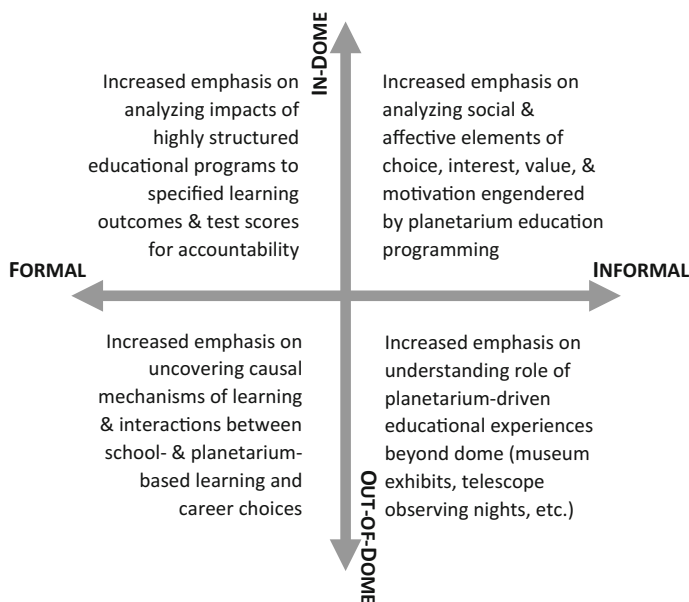
However it should be born in mind that these studies from over 40 years ago used a simplified 2D model of the universe projected on the dome. Detailed studies of the advantages of a well-planned classroom activity in a digital planetarium, with its greater ability to use spatial visualization and reinforce astronomical concepts through carefully use of visuals, is an area ripe for future research. One of the earliest attempts to study this employed a sample of students that had the same teacher over many years during which the switch from an optical to a digital planetarium occurred.

## Formal Setting

Formal learning environments are generally aimed at maximizing long-term content and cognitive gains for students, as well as developing long-term characteristics that encourage a growth mindset in learners. In order to evaluate these types of outcomes in formal environments, researchers can easily employ nearly any assessment tool including, but not limited to, pre-post-tests, interviews, course evaluations, observations, standardized tests, homework assignments, and case studies. This is usually much easier for the researcher, in that it is a controlled environment that can be altered to incorporate experimental teaching strategies, new curriculum, or activities. Also, the researchers can use two very distinct samples and track them for a longitudinal study. With an informal environment, it can be much more difficult to track participants over a long period of time due to changing phone numbers, emails, addresses, and other contact information. Additionally, planetarium education researchers could run into issues, such as sample bias, since participants in informal environments must partake in the research study voluntarily and be lead by their own motivation. In a formal environment, a sample can be very specific.

## Planetariums as a Formal Environment

Formal and informal learning environments, and the planetarium education research that occurs within them, do not always exist as being mutually exclusive, but instead sit along a continuum (Fig. 2.8) (Plummer et al. 2015). Planetarium education research studies often describe some blend of the two and how they interact with one another. For example, 3rd grade students learning about the lunar phases may take a field trip to the planetarium to develop a better understanding. Learning astronomical phenomena can take up a lot of working memory for a novice, and has been shown in research to be superior to the 2-D flat screen computer planetarium programs often used in the classroom or laboratory. Students are technically considered to be learning in an informal environment, but within a formal context. For the *international Studies of Astronomy Reasoning (iSTAR)* database, Tatge et al. (2016) and colleagues made specific judgments when categorizing studies as formal or informal. Using this particular example, this type of study would be classified as “formal” considering its context. However, say a study was being conducted on an astronomy class for adults held at a museum, which was taught within the planetarium. Technically, a museum is defined as an informal learning institution but a class could be considered more formal. Undoubtedly, the academic lines of distinction between formal education and informal education are unendingly blurred because planetariums in informal settings are often responsible for providing formal-like astronomy instruction to school students.



**Fig. 2.8** Quadrants of planetarium education research (adapted from Plummer (2015) and colleagues)

## Quantitative Methods

Regardless of whether a planetarium learning experience is specifically described as being a formal school-like learning experience or not, the rubber-meets-the-road task for the planetarium education researcher is most often the task of finding some rigorous and dependable way to scholarly describe, quantify, interpret, and explain causal mechanisms about cognitive and affective changes occurring in learners as a direct result of the planetarium learning experience. Planetarium education researchers most often use the time-tested and community accepted methods of empirical research.

Empirical research is generally defined as the scientific way of gaining knowledge through purposeful and systematic collection of evidence via objective observation, experience, and experiment. This type of research serves as the foundation for modern philosophy of science. In the present context, empiricism is a philosophical perspective on understanding the world that is in stark contrast to the differing perspectives of a priori reasoning, intuition, or revelation. Following that line of thinking, the basic empirical research tool kit of planetarium education researchers includes quantitative research methods, qualitative research methods, and mixed-research methods (Slater et al. 2015a).

Unquestionably, the most commonly used research methods in planetarium education research are those from the quantitative research realm. A comprehensive survey by Slater et al. (2016) and colleagues found that 65% of planetarium education research articles and dissertations relied solely on quantitative research methods, another 18% used a mixture of quantitative and other research methods, bringing the total to more than 80% of studies. This is likely a direct result of most planetarium education researchers having at least some formal academic background in the quantitative world of physical sciences like astronomy, which often unconsciously leans heavily on the Lord Kelvin's philosophical axioms that, "If you cannot measure it, you cannot improve it," and "when you cannot express it in numbers, your knowledge is of a meager and unsatisfactory kind".

In an attempt to quantify learning in the cognitive domain and attitudes in the affective domain, planetarium education researchers often use the broad notion of surveys to generate a numerical measure that describes a learner. There are numerous surveys, tests, conceptual diagnostics, attitude inventories, observation protocols, and field-note tabulation schemes available to planetarium education researchers; however, few are widely agreed upon within the planetarium education research community to make study-to-study comparisons easy to accomplish. As a result of too few agreed upon instruments, many planetarium education researchers decide to create their own surveys. Although it might seem easy to create a quick multiple-choice quiz for attendees exiting the planetarium, the accurate construction of a reliable and valid survey is an unexpectedly time-consuming, highly complex, and rigorous task that most planetarium education researchers are academically unprepared to undertake. As a result, the greatest weakness of most planetarium education research studies is the quality, reliability, and validity of the pre-test and

post-test survey instruments used. Novice researchers are encouraged to partner with experts for testing and measurement when it becomes necessary to create a new survey instrument.

## The *t*-Test

By and large, planetarium education researchers mostly use traditional measures of central tendency to describe their collected data: mean, median, and mode. This allows the mean scores between two measurements—pre-test to post-test gains or post-tests between two groups—to be compared. Across all studies Slater et al. (2016) and colleagues surveyed, the most common statistical tool used in planetarium education research is the *t*-test. The *t*-test is a numerical recipe to determine the likelihood that the difference in test scores between two groups is likely to have occurred by chance. For the *t*-test to work, statistical significance is determined by the size of the differences between the groups' mean scores, the sample size of each group, and the standard deviations of the group's scores. It is important to note that statistical *t*-tests comparing mean scores only work consistently when the scores of both groups have similar variance and have a range of scores that are normally distributed.

## Qualitative Methods

Qualitative research methods, in contrast, are fundamentally different than quantitative research methods in both their form and underlying philosophical position (Corbin and Strauss 2014; Plummer et al. 2015; Slater et al. 2015b). In recent years, the research method that was once known widely as qualitative research is now more often described as *interpretive inquiry*. Interpretive inquiry is, at its core, a systematic research method to understand how study-participants make sense of educational events and transformative experiences. Planetarium education researchers using this approach analyze interviews and focus group transcripts, make observations of participants' behaviors, analyzing patterns of discourse between learners, and interpret the artifacts of learning, such as student projects, student writings, and student drawings. Unfortunately, novice planetarium education researchers sometimes initially shy away from interpretive research methods because they naively think such approaches are more subjective than quantitative approaches. Experienced researchers understand that there is no truth to the notion that quantitative approaches are more objective and qualitative and interpretive approaches are more subjective. In fact, many Ph.D. dissertations in astronomy education research over the last decade have been mostly qualitative and interpretive in nature (Bailey and Lombardi 2015; Buxner 2015; Slater 2008, 2010).

According to Creswell (2012), the five most common genres of qualitative and interpretive research methods can be categorized as: biography, ethnography, phenomenology, case study, and grounded theory. In brief, biographical research is a method of uncovering the life story and broad influence of an individual. Ethnographic research, in contrast, is focused on conducting observations and interviews to describe and interpret the collective cultural characteristics and experiences of a group of individuals. Phenomenological research is a research method used to explain how individuals or groups understand, experience, and make sense of a shared experience, such as a one-time viewing of a solar eclipse. Case study research methods are used to develop an in-depth interpretation of a historical situation, using interviews or systematic analysis of archived documents. Finally, grounded theory research methods are used when precious little existing information is already known about an individual or group of individuals’ learning experiences. The emphasis of grounded theory research methods is to use field data to generate a theoretical explanation that can eventually be tested through experimentation.

These quantitative and qualitative/interpretive research methods are separate and apart from a third research method known as mixed-methods research. Mixed-methods research does not only mean that the planetarium education researcher does some quantitative work and some qualitative and interpretive work. Instead, mixed-methods research is a strategy in which evidence is pursued from multiple angles in order to triangulate conclusions and develop a more complete and accurate picture of the educational learning experience. When considering mixed-methods research, it is worth emphasizing that one common point of misunderstanding among novice planetarium education researchers is when written responses or recorded interview transcripts are collected using the methods of qualitative research, but then words and phrases are counted and numerically tallied to determine the frequency, range, and domain of responses—this strategy has taken initially qualitative research data and converted it into quantitative research,

**Table 2.1** Abridged list of journals publishing planetarium education research

Abridged list of journals publishing planetarium education research
International Journal of Science Education
Communicating Astronomy to the Public (IAU)
Journal of Astronomy & Earth Sciences Education
Journal of Geoscience Education
Latin American Journal of Astronomy Education
Journal of Research in Science Teaching
Physical Review—Physics Education Research
The Planetarian ( <i>International Planetarium Society</i> )
School Science and Mathematics
<i>Astronomy Education Review (no longer being published)</i>
<i>Journal and Review of Astronomy Education and Outreach (no longer being published)</i>

thereby no longer being qualitative or interpretive research. To be clear, this scenario is not mixed-methods, it would be quantitative methods because the data in this case was converted into quantitative evidence.

Forty years ago, scholarly journals were disinclined to publish anything but research articles solidly based on statistically rigorous quantitative methods. In today's modern era of planetarium education research, all respected journals will publish well-written, carefully done, and timely studies using quantitative, qualitative/interpretive, or mixed-methods research studies. Although there is a misperception among novice planetarium education researchers that only large sample size, multiple-group comparison studies are important and worth publishing, the truth is that that a well-executed case study of just a single person can have as much influence on the field as any other study. A non-exhaustive list of respected scholarly journals that publish planetarium education research is included in Table 2.1.

## References

- Akey, J. M. (1973). *The behavioral selection of planetarium concepts appropriate for second grade students*. Ed.D. Dissertation, University of Northern Colorado—Greely.
- Anderson, L. W., Krathwohl, D. R., & Bloom, B. S. (2001). *A taxonomy for learning, teaching, and assessing: A revision of bloom's taxonomy of educational objectives*. Allyn & Bacon.
- Bailey, J. M., & Lombardi, D. (2015). Blazing the trail for astronomy education research. *Journal of Astronomy & Earth Sciences Education*, 2(2), 77–88.
- Ball, D. L., & Forzani, F. M. (2007). Wallace foundation distinguished lecture—What makes education research “educational”? *Educational Researcher*, 36(9), 529–540. 2007.
- Buxner, S. R. (2015). Exploring how research experiences for teachers changes their understandings of the nature of science and scientific inquiry. *Journal of Astronomy & Earth Sciences Education*, 1(1), 53–68.
- Bloom, B. S. (1956). *Taxonomy of educational objectives. Vol. 1: Cognitive Domain*. New York: David McKay, 20–24.
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Chicago: Rand McNally.
- Corbin, J., & Strauss, A. (2014). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. Sage Publications.
- Creswell, J. W. (2012). *Qualitative inquiry and research design: Choosing among five approaches*. Sage.
- Chronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, 16(3), 297–334.
- Falk, J. H., & Dierking, L. D. (2012). *Museum experience revisited*. Left Coast Press.
- Falk, J. H., Storksdieck, M., & Dierking, L. D. (2007). Investigating public science interest and understanding: Evidence for the importance of free-choice learning. *Public Understanding of Science*, 16(4), 455–469.
- Feder, M. A., Shouse, A. W., Lewenstein, B., & Bell, P. (Eds.). (2009). *Learning science in informal environments: People, places, and pursuits*. Washington, DC: National Academies Press.
- Fisher, M. S. (1997). The effect of humor on learning in a planetarium. *Science Education*, 81(6), 703–713.

- Heyer, I., Slater, S. J., & Slater, T. F. (2013). Establishing the empirical relationship between non-science majoring undergraduate learners' spatial thinking skills and their conceptual astronomy knowledge. *Latin American Journal of Astronomy Education*, 16, 45–61. Retrieved from <http://web-02.ufscar.br/relea/index.php/relea/article/download/182/248>
- Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (1964). *Taxonomy of educational objectives: Vol. 2. Affective domain*. New York: David McKay.
- Novak, J. D. (1963). A preliminary statement on research in science education. *Journal of Research in Science Teaching*, 1(1), 3–9.
- Plummer, J. D., Schmoll, S., Yu, K. C., & Ghent, C. (2015). A guide to conducting educational research in the planetarium. *Planetarian*, 44(2), 8–24, 30.
- Reed, G. (1973). The planetarium versus the classroom—An inquiry into earlier implications. *School Science and Mathematics*, 73(7), 553–555.
- Ridky, R. W. (1974). *A study of planetarium effectiveness on student achievement, perceptions and retention*. Paper presented at the National Association for Research in Science Teaching, Chicago, IL. Available online at: <http://eric.ed.gov/?id=ED091207>
- Schleigh, S. P., Slater, S. J., Slater, T. F., & Stork, D. J. (2015). The new curriculum standards for astronomy in the United States. *Latin American Journal of Astronomy Education*, 20, 131–151. Retrieved from: <http://www.relea.ufscar.br/index.php/relea/article/viewFile/229/313>.
- Singer, S. R., Nielsen, N. R., & Schweingruber, H. A. (Eds.). (2012). *Discipline-based education research: Understanding and improving learning in undergraduate science and engineering*. Washington, DC: National Academies Press.
- Slater, S. J. (2010). *The educational function of an astronomy REU program as described by participating women*. Ph.D. Dissertation, University of Arizona.
- Slater, S. J. (2015). The development and validation of the Test Of Astronomy STandards (TOAST). *Journal of Astronomy & Earth Sciences Education*, 1(1), 1–22.
- Slater, S. J., Slater, T. F., Heyer, I., & Bailey, J. M. (2015a). *Conducting astronomy education research: An astronomers' guide, 2nd edition*. Hilo, Hawai'i: Pono Publishing.
- Slater, S. J., Slater, T. F., Heyer, I., & Bailey, J. M. (2015b). *Discipline-based education research: A scientists' guide, 2nd Edition*. Hilo, Hawai'i: Pono Publishing.
- Slater, T. F. (2008). The first big wave of astronomy education research dissertations and some directions for future research efforts. *Astronomy Education Review*, 7(1), 1–12.
- Slater, T. F., & Adams, J. P. (2003). *Learner-centered astronomy teaching: Strategies for ASTRO 101*. New Jersey: Prentice Hall.
- Slater, S. J., Tatge, C. B., Bretones, P. S., Slater, T. F., Schleigh, S. P., McKinnon, D., et al. (2016). iSTAR First light: Characterizing astronomy education research dissertations in the iSTAR database. *Journal of Astronomy and Earth Sciences Education*, 3(2), 125–140.



Research on Teaching Astronomy in the Planetarium

Slater, T.F.; Tatge, C.

2017, VI, 131 p. 28 illus., Softcover

ISBN: 978-3-319-57200-0