

Advancements in astronomy education research: Two decades of progress with undergraduate and adult learners

Molly N. Simon^{1,†} and Janelle M. Bailey^{2,*}

¹*School of Earth and Space Exploration, Arizona State University, Tempe, Arizona 85287, USA*

²*Department of Teaching and Learning, Temple University, Philadelphia, Pennsylvania 19122, USA*



(Received 3 September 2024; accepted 25 June 2025; published 5 August 2025)

In 2003, Bailey and Slater [A review of astronomy education research, *Astron. Educ. Rev.* **2**, 20 (2003)] published a review of astronomy education research (AER) to survey the field, identify areas of robust coverage, and uncover gaps in the research to help guide future researchers. While our review is not intended as a direct comparison to their original work, we aim to update the field's progress since that time. After initially considering more than 450 works, primarily peer-reviewed publications, we focused on research conducted in undergraduate settings or related to non-K-12-teaching adults, such as planetarium professionals, participants in massive open online courses, and citizen scientists. We identified seven broad categories of research, continuing several categories from the original study, such as student understanding and efficacy of instructional strategies, while also introducing new topics, such as authentic research experiences. These findings demonstrate the evolution of AER over the past two decades. The largest body of research remains focused on the efficacy and implementation of instructional strategies. Progress has also been made in exploring technology use (e.g., simulations) in introductory astronomy classrooms and integrating active learning into fully online courses. Notably, course-based undergraduate research experiences have emerged as a new focus in AER. Moving forward, future research should better incorporate educational theory (e.g., mechanisms of learning) and explore the unique educational opportunities offered within astronomy.

DOI: [10.1103/jp2p-97jb](https://doi.org/10.1103/jp2p-97jb)

I. INTRODUCTION AND PURPOSE

Modern research builds upon what came before, whether incrementally adding to existing lines of evidence or refuting prior work while offering alternative ideas. Knowing the field in which we work is, therefore, critical to moving that field forward. Literature reviews, whether narrative [1] or systematic [2], can provide guidance on the status of the field with greater or lesser detail depending upon the review's specific focus. Astronomy education research (AER) has benefited from such reviews in the past. Bailey and Slater [3] provided a broad overview of the AER landscape, documenting the main areas of focus and noting trends in publishing venues up to that point. The two decades since that review's publication have brought opportunities, growth, and challenges in the field. Accordingly, an updated review is now warranted.

*Contact author: janelle.bailey@temple.edu

†Contact author: mnsimon2@asu.edu

Published by the American Physical Society under the terms of the Creative Commons Attribution 4.0 International license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI.

II. BACKGROUND

A. AER in two worlds

AER typically can be found in two communities. The first—and older—is the broad science education research community. With publications from as early as the 1900s about students' understanding of concepts important to astronomy (e.g., heat and fire [4]) and the midcentury relating to specific astronomical objects (e.g., the Moon [5,6]), this research typically focuses on K-12 students or teachers, is likely conducted by science education researchers, and may be focused on a general construct (e.g., spatial thinking and learning progressions) within an astronomy context. An early review by Wall [7] included a total of 58 studies, 40 of which were at the elementary or secondary levels. Lelliott and Rollnick [8] presented a narrative review covering 103 articles over 35 years of AER focused on primary and secondary schooling, teacher preparation, and museum and informal settings; they purposefully excluded research at the postsecondary level other than teacher education. They found that 80% of the studies covered five “big ideas” of conceptions of Earth, gravity, the day-night cycle, the seasons, and the Earth-Sun-Moon system ([8], p. 1771). The authors noted that even though some of these areas are well researched,

the findings may not be translated into actionable practice by teachers.

AER at the postsecondary level has more commonly been the purview of discipline-based education research (DBER)—the second community. DBER is typically pursued by researchers for whom astronomy, as opposed to science education, is their primary or original field [9]. This work often occurs in the context of “ASTRO 101” courses: introductory or survey courses for undergraduates that cover the whole or a large chunk of the Universe and often satisfy general education science requirements for nonscience majors. Particularly in the early years, AER researchers may have had little or no formal training in education research or application (i.e., teaching). More recently, and much like in physics education research, graduate students have had the opportunity to focus on astronomy education research within their coursework, dissertation research topics, or both, although formal degree programs in AER are still few and far between.

B. Bailey and Slater’s review of 2003

The origins of Bailey and Slater’s [3] review were found in Bailey’s exploration of potential topics for her dissertation. Because AER was found across two fields and numerous journals, the review served as a way to explore what research had been conducted to date and to identify potential gaps. The project led to the inclusion of more than 100 articles, books, and gray literature (conference papers, dissertations, etc.) in three major sections:

1. Early reports on astronomy education (i.e., past reviews and “status” papers [7]).
2. Research on student and teacher understanding of various astronomy topics, including lunar phases, shape of Earth, day/night cycle, cosmology, astrobiology, and multitopic assessments.
3. Research on instructional methods or issues, such as personalized systems of instruction, conceptual change, collaborative learning, specific curricular programs, and planetarium use.

The present review is intended to provide a high-level overview of the evolution and current state of the field, with a specific focus on astronomy education research (as described in Sec. III A) and drawing parallels between methodologies and how they may have evolved over the past two decades. Additionally, we hope to identify areas in which AER stands apart—or can, in the future—from its discipline-based education research peers (e.g., physics education research). Whereas the 2003 review served as a model from which to structure and organize this manuscript, this work is not intended to be a direct comparison to Bailey and Slater’s [3] review. It differs from the 2003 review most notably in the omission of papers that discuss preservice or inservice teachers. The reason for this exclusion is that this review is intended to serve as the first in a two-paper series, the second of which will focus on

K-12 astronomy education research and will follow the model presented in [8], where K-12 students and teachers are discussed in the same manuscript.

III. METHODS

A. Article selection

Astronomy Education Review was the primary journal in astronomy education between 2001 and 2013 and was supported initially through NASA’s Science Mission Directorate, the National Optical Astronomy Observatory, and the Astronomical Society of the Pacific, and later by the American Astronomical Society. After *Astronomy Education Review*’s dissolution in 2013,¹ astronomy education researchers began publishing their work in a variety of other journals, predominantly *Physical Review Physics Education Research* (PRPER), the *Journal of Astronomy and Earth Sciences Education* (JAESE), and the *Journal and Review of Astronomy Education and Outreach* (JRAEO). In 2019, following the inaugural AstroEdu conference in Garching, Germany, members of the astronomy education community banded together to form the *Astronomy Education Journal* (AEJ). AEJ’s first issue was published online in November 2021.

To select articles for this review, we began by collecting and archiving all *Astronomy Education Review* articles published after the 2003 review. This was followed by a manual search of PRPER, JAESE, JRAEO, and AEJ before performing a more expansive Google Scholar search for astronomy education research articles published in journals such as the *International Journal of Science Education*, *Numeracy*, and the *American Journal of Physics*. Next, we compared the search results with our personal libraries (e.g., EndNote database, cited pieces in our own works) of previously identified AER articles. We also found new articles through ancestral searches of citations included in previously identified articles [10]. Finally, any new qualifying articles published after the main search process occurred (Winter 2023) were added when they became available. A total of 463 articles were included in a preliminary list. Once the preliminary list of articles was compiled, we worked meticulously to define a set of criteria to separate the articles into either “Include” or “Not Include.” Articles were classified as “Include” if they met the following criteria:

- The article was peer reviewed and published between the original 2003 review and December 2023.
- The article was a research article (e.g., there was an explicit research question and/or data collection,

¹At some point, AAS transferred the archiving of *Astronomy Education Review* to Portico, where all articles can currently be found. Unfortunately, as of the writing of this work, there has been a loss of access to supplemental materials associated with some articles.

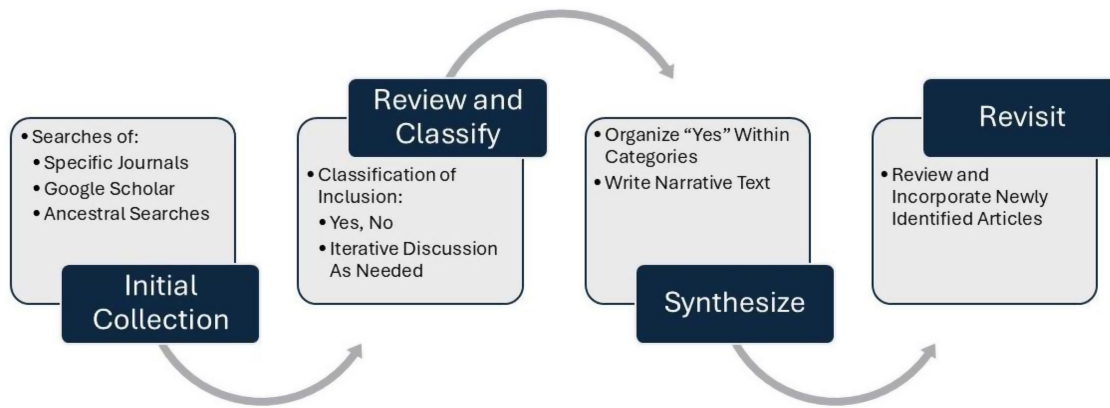


FIG. 1. A schematic representation of the process for the selection and categorization of articles in this review.

assessment validation, or classroom testing process) OR the article was a literature review.

- The article focused on undergraduate students and/or adult learners in the research population. Review papers had to include studies with these populations and could not focus on K-12 students or teachers exclusively.
- The article was published in English.

The articles selected for this review were refereed journal articles as well as one peer-reviewed conference proceeding review (Sec. IV A). Opinion pieces, press releases, dissertations, or theses without subsequent publications, and studies focusing on K-12 education, students, or teachers (inservice or preservice) were grouped as “Not Include.” Studies on K-12 education (with either students or teachers), in particular, have been saved for a potential future review.

Additionally, we identified 19 papers that we refer to as “How-To” articles. These How-To articles either (a) described how to conduct AER or some particular aspect of it or (b) discussed the implementation of research-based strategies into undergraduate teaching. Similarly, we identified many articles that were descriptions of new curricular materials without subsequent data collection or pilot testing efforts. We excluded How-To articles and curricular overviews from this review, ultimately choosing to focus exclusively on articles with a research emphasis (inclusion of research questions, classroom testing, and/or data collection and interpretation beyond course evaluations). Examples of articles that were omitted from this review include, but are not limited to, Birriel [11], Impey *et al.* [12], and Wallace *et al.* [13]. These example articles provide descriptions of demonstrations, educational resources, or classroom activities without subsequent research components. The final numbers of articles in the “Include” and “Not Include” groups were 135 and 328, respectively. A graphical representation of the article selection and iteration process can be found in Fig. 1.

B. Assignment of more specific categories

Articles from the “Include” group were originally sorted into one of eight categories: (1) Review Papers, (2) Studies on Student Understanding, (3) Assessment of Learning and Other Outcomes, (4) Efficacy of Active Learning Approaches, (5) Authentic Research Experiences, (6) Studies with Faculty and Planetarium Professionals, (7) Studies with Adult Learners More Broadly, and (8) Online Learning and Citizen Science. Within three of those eight categories, several more specific subcategories emerged. After three rounds of iteration, we placed articles into the category we believed best fit the main focus of the article, as a result of carefully reviewing each article’s purpose, research questions, data analysis, and conclusions. For example, Simon *et al.* [14] discuss the development and evaluation of a citizen science-based astronomy curriculum aimed at improving undergraduate students’ data literacy in online astronomy classrooms. Originally, that article was sorted into both the (8) Online Learning and Citizen Science and the (4) Efficacy of Active Learning Approaches categories. Upon more careful consideration, the article is truly a research study assessing the efficacy of a new piece of active learning curriculum (the citizen science and online components are secondary features). As such, this article was retained only in the Efficacy of Active Learning Approaches category and Active Learning in Online Courses subcategory. After completing this process with each of the 135 articles in the “Include” group, we dissolved the Online Learning and Citizen Science category, as each article within that category was a more complete fit somewhere else. The seven categories make up each of the subheadings in Results (Sec. IV) of this paper and are listed as categories 1–7 above. The final number of articles in each category is presented in Table I. Additionally, a complete list of the papers reviewed in this

TABLE I. Number of articles in each of the final categories for the “Include” group. Italics indicate subcategories. The sum of the italicized numbers is equal to the total number of articles within each category (in normal). The total number of articles reviewed in this work is in bold in the final row of the table.

Category	Number of articles
Review papers	12
Studies on Student Understanding	25
<i>Understanding of Astronomical Content</i>	23
<i>Understanding of Broader Aspects of Science</i>	2
Assessment of Learning and Other Outcomes	31
<i>Instrument Development for Astronomical Content</i>	23
<i>Assessments of Additional Variables</i>	3
<i>Assessing Affective Outcomes in Astronomy</i>	5
Efficacy of Active Learning Approaches	45
<i>The Big 3</i>	18
<i>Additional Active Learning Interventions in the In-Person Classroom</i>	14
<i>Active Learning in Laboratory Courses</i>	5
<i>Active Learning in Online Courses</i>	8
Authentic Research Experiences	3
Studies with Faculty and Planetarium Professionals	10
Studies with Adult Learners More Broadly	9
Total	135

work and their categorization is presented in Table II in the Appendix.

IV. RESULTS

Our results are divided into sections based on the seven categories described above. The ordering of the categories is not based upon frequency, chronology, or alphabet, but rather moves from the broadest category (reviews) toward student understanding, a primary category of the 2003 review (and others, e.g., [8]). A new category of assessment of learning is next, followed by the efficacy of instructional strategies, and then authentic research experiences. Most of the aforementioned categories are focused on undergraduate students in astronomy (often, though not exclusively, ASTRO 101), and the term “students” should be understood in this sense unless otherwise specified. The final two categories, which focus on adult learners, wrap up the presentation order.

A. Review papers

Despite AER being a relatively small field, a number of authors have taken on the task of writing literature reviews. These papers can equip researchers who are new to the field or a specific line of research with the lay of the land and often make recommendations for future research efforts. The reviews included here provide overviews of particular aspects of astronomy education publications as well as more focused reviews around particular topics.

Bretones and Megid Neto [15] looked at more than 230 papers presented in IAU proceedings between 1988 and 2006; none of the major categories they developed were

astronomy education research, although research papers were certainly included among other groupings such as “new trends in astronomy teaching” and “teaching and learning astronomy.” Fraknoi [16] discussed the history of where AER had been published, including a description of the rise and fall of *Astronomy Education Review*, whereas Bailey and Lombardi [9] described the history of individuals involved in AER and different foci between discipline-based education researchers and science education researchers. Bretones *et al.* [17] reviewed a decade of AER in the Brazil-based *Latin-American Journal of Astronomy Education*, which publishes articles in Portuguese, Spanish, or English (and abstracts in all three).

Content-focused reviews were also developed during this period. Kavanagh and colleagues wrote three reviews that were designed to guide curriculum designers and teachers by summarizing the existing research on students’ understanding of moon phases [18], gravitational free fall [19], and gravitational trajectories and orbits [20]. Madsen *et al.* [21] created a compendium of research-based assessment instruments that included astronomy as well as physics topics. Research on teaching in the planetarium was the focus of a review by Slater and Tatge [22]. Two reviews appeared in PRPER’s 2018 Focused Collection on AER: students’ perspectives on cosmological time [23] and spatial thinking in astronomy [24]. The opening chapter of peer-reviewed conference proceedings provided a review of research relating to robotic telescopes and astronomical research by students [25]. As the field grows, these more focused reviews provide insight into areas where there is a more robust research base. The present review, in contrast, is intended to take a higher-level view of the overall status of the field.

B. Studies on student understanding

The earliest papers identified by Bailey and Slater [3] involved investigations of students' understanding of particular astronomical content. Historically, this has been a common approach across science education [26]: identify what students know or don't know so that we can design or improve upon instructional strategies and curricular supports to facilitate greater learning. This focus has continued to be relevant in more recent years, although with some differences in both the content focus and the theoretical underpinnings. For the present study, we identified both AER investigations focused on particular astronomy topics and those targeting broader scientific understanding that are either directly relevant to astronomy or can use astronomy as a contextual vehicle for learning.

1. Understanding of astronomical content

Perhaps the largest category in Bailey and Slater's [3] review was research that cataloged students' understanding—or misunderstanding—of specific astronomy topics or content. At that time, a number of studies focused on geocentric topics, such as the causes of lunar phases or diurnal motion, or Earth's shape, all of which were consistent with studies being undertaken at the K-12 level (see Ref. [8] for a review of this area of research). Although these topics continued to be present in the subsequent years, there has been a shift toward focusing on concepts of interest at the ASTRO 101 level. During this time, there was also a movement in how such research was framed: There has been less focus on identifying “misconceptions” that need to be targeted or corrected and more acknowledgment that students enter the classroom with specific ideas that may help or hinder their learning processes (and that being aware of such ideas can support effective instruction).

Newer research around students' understanding of gravity has paid attention to ideas about Keplerian orbital motion and Newtonian gravity, both on Earth and within the solar system. Yu *et al.* [27] conducted interviews with more than 100 ASTRO 101 students in the first week of their course, focusing on Kepler's laws and how planets remain in orbit around the Sun. About three quarters of students believed planetary orbits to be highly elliptical (as opposed to having only slight eccentricities), but there were a number of ways in which their responses to questions were consistent with Kepler's second and third laws. A study using free-response questions and interviews to elicit students' ideas about gravity found some similarities to earlier studies as well as new ideas [28]. The authors proposed three emergent mental models of gravity and a fourth of “disparate misconceptions,” as well as three different ways in which students misapplied the scientific model in their understanding. This study formed the basis for instrument development as described in Sec. IV C 1.

The topic of stars is typically given a fair bit of time in many ASTRO 101 courses, but little had been investigated about students' understanding of them in the prior review [3]. Agan [29] interviewed high school students (eight from a first-year Earth science course and four from an upper-level astronomy course) and five undergraduate students with no astronomy background about their understanding of stars. The high school students tended to have more scientifically aligned responses than the other groups, such as identifying the Sun as our closest star and noting that a star has a unique energy production mechanism. Bailey *et al.* [30] used open-ended surveys with more than 2000 ASTRO 101 students to gauge their ideas about the nature, properties, and formation of stars. Thematic coding of the results—a mix of scientifically consistent and other ideas—contributed to the development of a multiple-choice instrument (see Sec. IV C 1). Constellations and the names of bright stars were the focus of work by Hintz *et al.* [31] conducted in association with use of a planetarium, with preinstructional assessments of more than 650 undergraduate students (and a smaller number of early high school students). Students self-reported that they would be able to identify an average of two constellations (typically Orion and Ursa Major) and about the same number of stars (with Polaris and Betelgeuse being most common) prior to instruction.

Uncovering students' ideas around cosmology was a dominant topic of this period. Using a suite of open-ended surveys prior to instruction, Bailey *et al.* [32] looked at multiple cosmological and related foundational ideas, such as the definition of a light year; composition and structure of the Universe; evolution of the Universe; and the nature of dark matter and dark energy. This work was expanded through the use of similar surveys plus in-class assignments, exam questions, and interviews to provide deeper investigations into cosmological distances and structure [33], the composition of the Universe [34], the Big Bang model [35], fate of the Universe [36], and curvature of the Universe [37]. Around the same time, Wallace *et al.* published a suite of studies on cosmology education, including one that explored students' preinstructional ideas about “expansion and evolution of the universe, the Big Bang, interpreting Hubble plots, and the evidence for dark matter in spiral galaxies” [38] (pp. 010104–1; see also Secs. IV C 1 and IV D 1 for related studies from this team). In all these cases, the researchers identified areas of both consistency and inconsistency with scientific understanding of the cosmological content.

Eriksson *et al.* [39] surveyed students and astronomy instructors (graduate teaching assistants and professors) to see what they discerned (defined as “noticing, reflecting, and constructing meaning,” p. 19) about three-dimensionality when viewing a simulation of a fly-through of the Milky Way and onto other galaxies. The researchers identified six categories of discernment, ranging from

noticing motions (reflecting one-dimensional perspectives) to advanced three-dimensional noticings; the latter category was more often noticed among professors and graduate students than advanced or early undergraduates, as might be expected.

Buck [40] randomly assigned students to one of three groups to observe a visualization representing dark matter over time, each using a different color scheme. She found that there was a difference in how the groups interpreted the visualization (i.e., identifying where stars, hydrogen gas, dark matter, or empty space are located) based upon the colors used. Buck Bracey [41] used card sorting and drawing activities, finding a plethora of cosmological ideas that “superficially appears to reveal that a group of community college students have failed to replace a number of pervasive and basic misconceptions about the Universe” (p. 941). However, she also noted the commonalities that many of these ideas had to historical ideas about cosmology, as well as connections to spiritual beliefs, and presented multiple instructional implications for viewing these ideas not as misconceptions but as building blocks to support learning in progress.

Gross and Lopez [42] found through a pretest/posttest plus interview study that upper-division undergraduate and graduate students struggled to understand representations of solar wind, often confusing it with the planetary magnetic field. This difficulty is connected to common presentations of plasmas that students may apply in a way that is inconsistent with the intent or appropriate context. Simon *et al.* [43] surveyed more than 1000 students prior to instruction about their ideas on planetary formation. Many students initially believed that our solar system was created by the Big Bang (similar to results found in other work [32,35]) and were limited in their knowledge of foundational ideas that would support robust understanding in this area. Bailey *et al.* [44] used open-ended questions relating to space exploration postinstruction and found that students shared incomplete but reasonable ideas about the factors that both crewed and uncrewed missions must take into consideration, such as fuel type, hazards, and travel time.

The various studies up to this point have focused on a single topic (or perhaps multiple closely related concepts), whereas a few additional papers report on multitopic investigations. Zeilik and Morris [45] used questions from the Astronomy Diagnostic Test (ADT; [46]) as part of a larger instrument to investigate pre-to-post-test change for an astronomy course for science, technology, engineering, and mathematics (STEM) majors, finding notable overlapping areas of difficulty with the traditional ASTRO 101 population as well as other common challenging concepts previously identified in physics (e.g., inverse square law challenges). They also administered a 34-item “attitudes” survey (encompassing questions relating to affect, cognitive competence, value, and difficulty as each relates to

astronomy specifically and science more generally) pre- and post-test and found no significant changes, similar to findings in previous studies by Zeilik *et al.* [47,48]. Simonelli and Pilachowski’s [49] study focused on five different areas via open-ended questions with ASTRO 101 students: relative scales, arrangements, and abundance of objects in the Universe; different types of objects; origins of the solar system; differences in characteristics between planets and stars; and issues around life in the Universe. The authors found a mix of ideas, including information both accurate and not, with the formation of the solar system often confused with the formation of the Universe (i.e., Big Bang), among other findings. Heyer *et al.* [50] found that although students made small gains over the course of the semester on two different instruments, their pre-to-posttest gain scores were correlated with their spatial thinking ability as measured by a third instrument designed for that study. Slater *et al.* [51] dove into students’ understanding via a detailed analysis of each question in a multiple choice instrument (developed by Slater [52] and described in Sec. IV C 1). They found that difficulties on individual items tended to align with other investigations for those that exist, although detailed studies were not present for every topic on the assessment. Using the Conceptual Frameworks in Astronomy (CFA) instrument, Pundak *et al.* [53] classified undergraduates from three colleges as having one of or some combination of pre-scientific, geocentric, heliocentric, or sidereal (scientific) mental models. They found that most students held a mix of models and that the model held may vary depending on the topic.

2. Understanding of broader aspects of science

A final area of work on students’ understanding relates to the broader issues of the nature of science or scientific inquiry, wherein astronomy is used as the context of interest. Sugarman *et al.* [54] analyzed surveys from more than 1000 students on their ideas about astrology, finding that more than three quarters of the respondents indicated that astrology is “very” or “sort of” scientific (p. 4). This result was higher than national public surveys, although the reason for said differences was not clear. In a second astronomy course, undergraduates were asked to write two essays about whether or not sky observers from ancient civilizations would be considered scientists and a third about the extent to which they would be thought of as “good scientists” [55]. Students’ responses to these prompts demonstrated improved understanding of science over the course, through the use of more frequent and more advanced language such as discussing process, hypotheses, or data.

C. Assessment of learning and other outcomes

The last 20 years have seen the development of a number of assessment tools and strategies that would allow

researchers and educators to measure student understanding, learning over a course or intervention, or other characteristics. These have typically been developed in and for the ASTRO 101 context [56]. Articles in this category describe the development and validation processes for such instruments and often use classical test theory for statistical analysis. We identified three subcategories of instrument development: astronomical content, relevant other content, and affective outcomes.

1. Instrument development for astronomical content

Among the most common approaches for assessment in astronomy education has been the development of a *concept inventory*, typically a multiple-choice survey focusing on a single topic or a small number of closely related topics [57]. The Force Concept Inventory (FCI; [58,59]) in physics is perhaps the most widely known concept inventory, and the Lunar Phases Concept Inventory (LPCI; [60]) was the first to be developed in astronomy.

The Star Properties Concept Inventory (SPCI; [61]) addressed the nature, characteristics, and formation of stars. This instrument has been validated using both classical test theory [62] and item response theory [63], with the latter study providing broad insights into the applicability of this analysis approach to AER. Bardar *et al.* [64] described the development process for the Light and Spectroscopy Concept Inventory (LSCI), focusing on the nature of light and how astronomers use light and spectra to glean information about the Universe. Schlingman *et al.* [65] provided an analysis of a large dataset of LSCI responses from various ASTRO 101 courses across the U.S., and Wallace *et al.* [66] applied item response theory to the LSCI national dataset. Williamson *et al.* [67] created the Newtonian Gravity Concept Inventory (NGCI) for use in astronomy. The Size, Scale, and Structure Concept Inventory (S3CI) focused on common student difficulties when considering astronomical issues of various scales [68]. The Planet Formation Concept Inventory (PFCI; [69]) considered the processes and characteristics of planetary formation both within our solar system and in other stellar systems (i.e., exoplanets).

Other approaches to assessment focused on specific topics, although they may not have followed the norms for concept inventories [57]. LoPresto and Hubble-Zdanowski [70] created an assessment centered around the ideas in the Drake Equation, tested in both a “life in the Universe” honors seminar and a general ASTRO 101 course. The former showed greater pre-to-post instructional gains, as might be expected by the greater emphasis on the topics in that particular course. Wallace and colleagues created [71] and tested via both classical test theory [72] and item response theory [73] four open-ended surveys covering aspects of cosmology, including Hubble plots, models of the Universe, evolution of the Universe, and dark matter. These surveys were used to measure learning in

later studies (see Sec. IV D 1). Bekaert *et al.* [74] designed and validated with both high school and undergraduate students an instrument to gauge understanding of the apparent motion of the Sun and stars (AMoSS).

Another set of assessment measures focused on a broader array of content than those described so far. Favia and colleagues described the creation and internal validity [75] of the Astronomy Beliefs Inventory, and then showed how that instrument is correlated with other measures of content knowledge [76]. The instrument presented over 200 declarative statements in two versions—one in which every statement was scientifically inaccurate and the other in which there was a mix of accurate and inaccurate statements. Students then indicate their belief with the statements (i.e., in “true or false” fashion). The Test Of Astronomy STandards (TOAST; [52]) was an attempt to fully cover the range of astronomy content in the typical ASTRO 101 course, using items from existing instruments (e.g., ADT [46]; SPCI [62]; NGCI [67]; etc.) Pundak [77] described a new version of the Conceptual Frameworks in Astronomy instrument (the original was developed in the late 1980s), designed to distinguish whether students’ understanding could be considered prescientific, geocentric, heliocentric, or sidereal/scientific. This instrument was tested with engineering students, with results compared to the earlier version used with junior and senior high school students.

A secondary group of studies looked at previously developed instruments and their applicability for measuring change over time or comparing instructional strategies. Here, the focus was on to what extent the instrument could detect differences, as opposed to what those differences might be (which is sometimes the focus of articles in Secs. IV D 1 and IV D 2). LoPresto [78] used the ADT over three semesters at a community college to identify areas of greater or lesser gain, pre-to-post-test, (a) noting that the largest gains aligned with the major course goals and where greater instructional effort was put forth and (b) suggesting that the ADT was therefore a reasonable measure to use. Brogt *et al.* [79] found that there was no significant difference in normalized gain between courses identified as lecture only, lecture plus lab, lecture plus discussion, or lecture plus lab and discussion. This could mean that the broad content coverage of the ADT (which, by design, covers a wide number of topics, in contrast with concept inventories as described above) is not effective in measuring instructional differences. In contrast, the LSCI did show sufficient sensitivity in differentiating learning gains between traditional and active learning courses from pre-to post-test [80].

In addition to instruments to measure content knowledge or learning, at least one instrument has been created to assess affective outcomes. Freed and colleagues developed a measure of self-efficacy for learning astronomy content and using astronomical instrumentation (primarily robotic telescopes; [81]). Two scales comprise the assessment—one on “personal self-efficacy” and the other on “astronomy

instrument self-efficacy.” Although the original instrument showed high reliability and construct validity for both scales [81], a subsequent confirmatory factor analysis demonstrated skewness in the original astronomy instrument self-efficacy scale, leading to modifications that yielded better outcomes [82].

2. Assessments of additional variables

Some of the instruments described here have been used to consider additional factors that might influence achievement. Wooten *et al.* [83] compared students’ performance on multiple-choice and open-ended questions, using a rubric to score the latter, with a particular eye toward students’ ability to provide evidence for their reasoning, measured at three time points. Students performed better on the multiple-choice questions, leading the authors to conclude that such instruments may lead to an overestimate of abilities, particularly at the end of a course. Robertson and Finch [84] looked at 10 years’ worth of student evaluation data in ASTRO 101 courses. They found that the highest correlation existed between what they termed “grade inflation” (i.e., “the difference between the grade expected in this course, as reported by the student, and the estimated GPA [grade point average] as reported by the student” (p. 32) and course evaluations. The authors also found positive correlations between estimated GPA and expected course grades; negative correlations between expected course grade and study time for the course; and that female students expected lower course grades than male students did. Negative correlations were found between evaluations and both background in physical science and student study time. Finally, Richards and Kelly [85] used a logistic regression model to determine predictors of performance in community college astronomy coursework. They found that predictors included mathematics achievement, remedial mathematics course enrollment, and enrollment in multiple astronomy courses, whereas traditional demographic variables (ethnicity, socioeconomic status, gender, age) were not significant predictors of astronomy performance.

3. Assessing affective outcomes in astronomy

Several studies have investigated connections between variables (e.g., knowledge) as measured by some of the previous instruments and affective outcomes, such as self-efficacy, interest, beliefs, and persistence in science. Bailey *et al.* [86] used a mixed methods design to identify changes in and relationships between students’ knowledge as measured on the SPCI, their interest in learning about stars, and their self-efficacy for learning about stars. They found that interest, which was relatively high to begin with, had neither significant change over the semester nor significant relations with knowledge change, but that self-efficacy was related to knowledge and increased in some of the courses while decreasing in others. Follow-up

interviews suggested that there were differences in sources of self-efficacy (particularly verbal persuasion and mastery experiences) as facilitated by the instructors of the participating courses, with those courses that provided greater verbal persuasion and opportunities for mastery experiences resulting in higher self-efficacy. Berryhill and Slater [87] conducted interviews with 11 students who scored either lower or higher than typical on the TOAST. After inquiring about a number of background factors such as academic performance, science coursework, first-generation college student status, family members working in STEM fields, etc., students were asked about their reasoning behind questions where their answer may have differed from the majority of the class. Although no formal statistics were presented, the authors found no suggested relationships between any of their hypothesized predictor variables and TOAST pretest scores.

In a 27-year-long study of more than 12 000 ASTRO 101 students, Buxner *et al.* [88] analyzed surveys of science knowledge, beliefs, and attitudes. The knowledge scores were compared to national samples of similar questions, while relationships between the three variables were also probed. Although the participants consistently scored higher than public sampling on similar questions, the trend over time suggests little change. Demographic variables and beliefs in and attitudes toward science accounted for 11% of the variance in knowledge scores. The authors also found that respondents who reported use of academic sources of information (e.g., textbooks and instructor) had significantly higher knowledge scores than those who relied on more general internet searches. Wallace *et al.* [89] found that perceptions of science and a number of aspects of science, as measured by the Thinking About Science Survey Instrument [90], for ASTRO 101 students were generally positive and improved slightly over the course. Bergstrom *et al.* [91] surveyed nearly 16,000 first-year English students across 27 U.S. institutions about their interest in various STEM fields as a career and how it changed over time from middle grades to the beginning of college (per their reflections). For astronomy, the average interest for both female and male respondents was higher than the average of other STEM fields at middle school, but lower at three later time points, and decreased from the first to the last time. Many students who indicated an interest in astronomy moved into other STEM fields over time. Female respondents indicated lower interest in astronomy than male respondents at nearly all points.

D. Efficacy of active learning approaches

Over the last 20 years, there has been a tremendous increase in the number of published studies focusing specifically on the impacts of active learning strategies on student comprehension, especially as compared to traditional lecture. While there is not a singular, universally adopted definition of active learning, it is often considered a

student-centered instructional approach where students help facilitate their own learning. In collaboration with expert teams of astronomy, biology, chemistry, engineering, geography, geoscience, and physics education researchers, Lombardi *et al.* [92] defined active learning in the STEM classroom as:

...a situation in which the instructor and instructional activities explicitly afford students agency for their learning. In undergraduate STEM instruction, it involves increased levels of engagement with (a) direct experiences of phenomena, (b) scientific data providing evidence about phenomena, (c) scientific models that serve as representations of phenomena, and (d) domain-specific practices that guide the scientific interpretation of direction observations, analysis of data, and construction and application of models (p. 16).

When implemented successfully, active learning strategies have the ability to improve exam scores [93], student understanding and content comprehension [94], and may help to bridge the gap in student apprehension between different demographic groups [95,96]. Due to these aforementioned benefits, many studies have attempted to assess the impact of active learning in a variety of undergraduate astronomy course contexts.

1. “The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share

One of the most studied active learning strategies in the introductory astronomy classroom context is the *Lecture-Tutorial* (LT). LTs are short (15–20 minutes), pen-and-paper activities that include carefully sequenced Socratic-style questions aimed at addressing common conceptual and reasoning difficulties about specific topics [e.g., [94,97,98]]. LTs are designed to supplement traditional lectures and are to be completed collaboratively by students working in groups of 2–3. In the inaugural LT study, Prather *et al.* [97] used a 68-item concept inventory to assess the difference in ASTRO 101 students’ learning at two time points. At the completion of the day’s lecture, students were given 3–4 multiple-choice items from the 68-item assessment to answer. They next spent an additional ~15 minutes completing the corresponding LT and were then again presented with the same 3–4 questions. Most notably, the authors found a 20% improvement in students’ mean assessment scores after completion of the LT when compared to lecture alone.

A study by Alexander [99] used a modified version of the ADT [46] to compare traditional, mostly lecture-based introductory astronomy courses with two updated versions of the course that included ten LTs from the inaugural *Lecture-Tutorials for Introductory Astronomy* workbook

[100]. The ADT was administered at the beginning and the end of the course, rather than immediately after the LT as done by Prather *et al.* [97]. Alexander [99] found no statistical change in students’ normalized gain scores between the updated, more interactive versions of the course and the lecture-based offerings, but noted it was still worthwhile implementing teaching strategies in the updated courses that encouraged students to more actively engage with the material.

LoPresto and Murrell [101] used the SPCI to compare student learning gains between two sections of introductory astronomy that included LTs when teaching about stars, and two sections of the same course that taught stars using traditional lecture. In this study, students in the sections that implemented tutorials experienced higher learning gains ($g = 0.50$) than the traditional lecture course ($g = 0.41$). These results were more consistent with those found in the Prather *et al.* [97] study, although the learning gain differences in the LoPresto and Murrell study between the more interactive course sections and the lecture-based sections were not quite as vast. Eckenrode *et al.* [102] analyzed the quality of ASTRO 101 students’ written responses to specific LT questions to determine if the quality of their responses was correlated to LSCI and class exam scores. Although the authors were unable to find a correlation between written response quality and assessment scores, when comparing students’ normalized gain scores to a broader population of ASTRO 101 students, the students in the courses where LT written responses were analyzed had normalized gain scores within the top 10% of the nation [103]. These findings lent support to the claim that LTs bolster student understanding more broadly.

Although preliminary studies of LTs yielded mixed results, several researchers developed tutorials on additional introductory astronomy topics and tested their efficacy in a variety of classroom contexts. LoPresto [104] introduced the development of two visual-based LTs on the topics of comets and extrasolar planets. Students who completed the tutorials performed significantly better on a visual-based assessment (that included drawing and labeling) than those who received lecture-based instruction or visual guides separate from the tutorial. Wallace *et al.* [105] developed a suite of five cosmology LTs on the following topics: (1) the Big Bang, (2) dark matter, (3) Hubble’s Law, (4) the Universe and expansion, and (5) expansion, lookback time, and distances. These LTs were implemented with thousands of introductory astronomy students between 2009 and 2010, and their efficacy was assessed using one of four possible assessment forms that covered content related to the cosmology tutorials. Overall, the authors found that students in courses who implemented the cosmology LTs, as well as students in courses who did not implement the tutorials, saw gains in their conceptual understanding of relevant cosmology topics. Students in the LT courses, however, performed

better (although not always significantly) on all four survey forms than students who were in courses that did not implement the tutorials. Perhaps most interestingly, however, was that the authors asserted that learning gain variations amongst courses that used the LTs were likely attributed to different types of instructor implementation, which was not standardized throughout the study [105]. Because the authors did not observe how each instructor implemented the tutorials in their course, they could not determine whether each instructor's method engaged students sufficiently in the tutorial process. Wallace *et al.* [106,107] went on to develop two additional LTs on the topics of (1) molecular excitations and synchrotron radiation and (2) interferometry, respectively. The pilot testing of these tutorials in a variety of ASTRO 101 courses yielded similar results to Prather *et al.* [97], further highlighting the potential value and versatility of a 15-minute intervention specifically designed to promote a more active approach to learning.

LoPresto and Slater [108] evaluated the efficacy of LTs developed to teach solar system formation and planetary surfaces. Introductory astronomy students ($N = 144$) were given a set of multiple-choice questions on the topics of solar system formation and planetary surfaces as a pre-test. A smaller group of those students ($n = 49$) was given a relevant subset of questions again after a traditional lecture on the topic(s). Another group of students ($n = 72$) received traditional lectures and completed the corresponding LT before retaking the assessment. For both topics (solar system formation and planetary surfaces), students who completed the LT performed significantly higher on each individual assessment item compared to the students who received the standalone lecture. Additionally, the authors explored the impact of the visual-based tutorials described in LoPresto [104] on students' performance on related visual assessments on the topics of comets ($N = 251$ students) and exoplanet detection ($N = 284$ students). For both topics, the subset of students who received a lecture and completed the relevant visual tutorials scored significantly better on the corresponding visual assessments. Their results indicated that increases in student content comprehension were not solely due to the traditional LT approach, but could be similarly achieved using slightly modified visual-based LTs and shorter, diagram-based visual assessments [108].

Similar to LTs, Ranking Tasks (RTs) are 15–20-minute active engagement activities intended to supplement traditional lectures. First described in the physics education literature [109], RTs typically present students with images or diagrams that require learners to “make a comparative judgment and to identify the order or ranking of the various situations based on some physical outcome or result” [110, p. 3]. One example of an RT could involve placing a satellite at several different distances between Earth and the Moon in a superficial Earth-Moon system. Given the

masses of the objects relative to each other, students could be asked to rank the positions where the satellite experiences the largest to smallest gravitational force from one of the bodies in the system.

To assess the potential impact of RTs on student learning, Hudgins *et al.* [110] used a 28-item multiple-choice test measuring student understanding of eight topics covered in an introductory astronomy course ($N = 253$ students). The full assessment was administered on the first day of class, prior to any relevant instruction. Each time one of the eight relevant topics was covered in class, a subset of questions from the full assessment was administered at the culmination of a 30-minute lecture. This question subset was administered again after students' completion of 3–4 relevant RTs. Akin to the findings discussed in Prather *et al.* [97], assessment results indicated higher average assessment scores (77%) after completion of the RTs when compared to traditional lecture (61%), indicating that RTs, when implemented successfully, have the potential to significantly improve student understanding of an array of astronomical topics.

Think-Pair-Share (TPS), often referred to as peer instruction (PI; [111]), is an active learning strategy that presents students with an in-class (usually multiple-choice) question, requires them to *think* before voting on their answer choice, and then *pair* up with their surrounding peers to discuss their answers before they *share* by submitting a second response [112]. The goal of TPS questions is to have students use peer interaction as a tool to address potential conceptual and reasoning difficulties that may arise when introduced to new topics. For example, Crider [113] implemented a strategy intended to complement TPS questions titled “Hot Seat” questioning. This approach involved groups of four students sitting in the front of the classroom and working collaboratively to answer questions relevant to a specific day's lecture material. Students were given their Hot Seat dates in advance and spent extra time learning the lecture material as a result. Results from both midterm exams and the ADT indicated that students ($N = 50$) were more likely to answer questions correctly on topics aligned with their specific Hot Seat dates, suggesting a more involved version of TPS questioning could provide additional benefits to student learning.

Electronic clickers gained popularity in ASTRO 101 courses in the early 2000s as a tool to assess different strategies aimed at encouraging peer instruction and engagement when answering in-class questions. Len [114] found that students were more incentivized to work collaboratively (using a TPS-like model) when answering clicker questions if the class was rewarded with extra participation points when a certain course success threshold was reached. When students were not provided with the same success rate incentive, one-third of students chose to answer clicker questions independently, rather than working with their peers. Regardless of whether students

generally preferred to answer clicker questions independently or collaboratively, when students were incentivized to answer questions collaboratively, the overall course success rate was much higher ($\sim 80\%$). James *et al.* [115] recorded small PI groups over two semesters of an ASTRO 101 course to determine effective strategies for peer discourse when answering TPS-style clicker questions. Their findings suggested that when a high-stakes grading incentive was provided (little credit given for incorrect answers), small groups were typically dominated by one student confident in their answer choice. In these situations, small groups ultimately converged on a single answer choice. In low-stakes classrooms where TPS-style questions were ultimately graded for participation (given 90% of the credit just for answering), students tended to select an answer choice based on their current understanding (rather than acquiescing to the dominant student in the group). As such, the distribution of answers in low-stakes settings “may [have been] more representative of the degree of understanding that actually exist[ed] in a class” [115], p. 42].

Prather and Brissenden [116] determined that electronic clickers were as effective as ScantronTM forms in terms of gathering student data, but argued for the increased pedagogical benefits of analog, colored voting cards in lieu of electronic clickers when answering TPS questions. Regardless of the voting mechanism (electronic clickers or colored cards), previous studies conducted predominantly in physics education research indicate that TPS/PI questioning is a valuable active learning strategy for improving student understanding on a variety of introductory topics [e.g., [117]].

Prather *et al.* [103] assessed the efficacy of interactive instruction more generally by examining gain scores on the LSCI across 69 introductory astronomy courses at 31 different institutions. In addition to administering the LSCI to students, the authors provided instructors with an Interactivity Assessment Instrument (IAI) to determine the extent to which instructors implemented LTs, RTs, and TPS questions in their courses. Normalized gain scores varied across courses, instructors, and institution types (e.g., 2-year versus 4-year colleges), with the highest gains often observed in courses with high IAI scores. Not all interactive courses resulted in high LSCI gains, however, suggesting that *how* active learning strategies were incorporated into the classroom was an important contributor to student learning [94,103]. Rudolph *et al.* [96] built on the work of Prather *et al.* [103] by performing a demographic analysis of the students in those same courses. Their findings suggested that active learning strategies benefited a diverse population of students equally, including men and women, students from different ethnic backgrounds, and students with a variety of incoming mathematics backgrounds, primary languages, and incoming GPAs [96]. These results highlight the importance of implementing active learning strategies when instructors endeavor to create a more equitable learning environment.

Rudolph [118] found that LTs, RTs, and TPS questions, when integrated consistently throughout an ASTRO 101 course, improved student learning to the extent that a course textbook was no longer required. Despite using a variety of assessment methods (course test scores, LSCI, and SPCI), Rudolph [118] found no statistically significant difference between students’ assessment scores during semesters where a course textbook was required versus when the course textbook was made optional. Furthermore, in textbook-optional courses, assessment scores remained statistically indistinguishable between students who chose to purchase the textbook and students who did not. Clicker-type survey questions administered to students enrolled in the textbook optional version of the course across four different semesters indicated that the majority of students (67%) felt the textbook was not essential in a classroom setting that successfully implemented LTs, RTs, and TPS questions throughout the course.

Several of the active learning studies previously discussed looked at the effects of LTs, RTs, and/or TPS questions on student attitudes in addition to content comprehension. For example, Prather *et al.* [97] administered a 34-item Likert-style attitudinal survey [119] along with their 68-item multiple-choice assessment. Although students did not report statistically significant attitudinal changes after completion of their LT-based introductory astronomy course, whole-class focus group interviews indicated that students thought LTs were one of the most important components of the course, and the integration of LTs made the course more enjoyable overall. Similarly, Hudgins *et al.* [110] administered a Likert-style attitudinal survey of their own design to better understand students’ attitudes toward using RTs in their introductory astronomy courses. Of the 132 students who completed the survey, “83% of students believed that ranking tasks contributed positively to their learning, and 72% thought that the exercises helped them on tests” (p. 10). Sixteen percent of students reported negative attitudes toward RTs and preferred more passive types of course instruction (e.g., lecture only).

Prather and Brissenden [116] administered a 19-item attitudinal survey to determine the efficacy of clickers as a data gathering tool in a course that implemented both clicker-based TPS questions and LTs. While the majority of students “reported that the use of [clickers] contributed positively to their learning of course topics, achievement on exams, [and] interest in course topics,” students primarily credited LTs with helping them best prepare for exams and for reinforcing the course content (p. 5). LoPresto and Slater [108] gave students ($N = 264$) in their active learning-based course a Likert-style attitudinal survey to assess students’ perceptions about active instruction (including LTs, visual-based LTs, and TPS questions) versus traditional lecture more broadly. While the students did not express a clear preference for active instruction over

traditional lecture, their responses to an open-ended question on the survey indicated that students believed the active learning activities did help them learn more deeply.

2. Additional active learning interventions in the in-person classroom

Studies of the benefits of active learning strategies in the ASTRO 101 classroom are not limited to LTs, RTs, and TPS-style questioning. Astronomy education research over the past two decades has resulted in the evaluation of a variety of course-based interventions, curricular products, and innovative teaching strategies aimed at promoting more comprehensive student engagement that, in some cases, may lead to increased motivation as well. In one such example, Straits and Wilke [120] analyzed the effectiveness of a broader “activities-based” approach to teaching introductory astronomy and compared student performance on the ADT [46] in the activities-based course to a lecture-based control group. The instructors of the intervention course predominantly used activities from *Cooperative Learning Activities in Introductory Astronomy for Non-science Majors* [121], implementing 3–4 activities during each class period to supplement a short lecturette. While the authors did find statistically significant learning gains on students’ ADT scores in both the activities-based course and the control group, the scores of the students in the activities-based course were statistically indistinguishable from those in the control group. Responses from student interviews suggested that students benefited the most from activities that “helped [them] visualize spatial relationships, provided equal opportunity for engagement, and were clearly related to course content” [120], p. 46].

An area of study in particular has been on the successes and shortcomings of collaborative course assessments (e.g., quizzes and exams). Zeilik and Morris [122] introduced the concept of cooperative quizzes across two semesters of an ASTRO 101 course. The cooperative quiz format involved students completing a multiple-choice quiz before discussing the quiz questions and answers with a small group of their peers. After the discussion, the students had the opportunity to change their original answers to reflect their conversation with their peers. The quiz score ultimately became a weighted average of their individual scores and then their (potentially) modified scores after peer discussion. Unlike the results from James *et al.* [115], Zeilik and Morris [122] did not find that individual students dominated the peer group discussions. They instead found that the cooperative quiz approach resulted in average normalized gains of ~ 0.4 across the two semesters compared to students’ quiz scores before the collaborative component. They tested student retention of the course material (approximately one month later) by looking at a subset of repeat questions on course exams (the course exams were not collaborative). Students’ responses to repeated questions on the exams were statistically indistinguishable

from their responses to identical questions on the cooperative quizzes, indicating they retained the material between the quizzes and the course exams.

Miller and James [123] employed a similar approach by analyzing the effects of collaborative exams on ASTRO 101 students’ learning and attitudes. Students across two semesters took three to four exams that included two stages. In the first stage, students answered exam questions individually. After the individual stage, they were permitted to discuss a subset of exam questions with their pre-established small peer groups. Mid-semester, the students were given an assessment so instructors could better understand their attitudes toward collaborative testing and its impact on exam preparation. At the culmination of the semester, the students took a final exam without a collaborative component; however, many of the questions they had previously discussed collaboratively reappeared on the final exam. The researchers found higher normalized gain scores for the subset of final exam questions that students had been previously able to collaborate on. Overall, many students (72.8%) agreed that the collaborative component of exams helped them better understand the course material, and many students mentioned being encouraged to form study groups to prepare for the exams. Several students conversely noted they did not like that their group had to converge on one specific answer to move forward on the collaborative component of the exam. Overall, group dynamics remain complicated, but these studies indicate there may be benefits to collaboration within high-stakes testing environments.

Another strategy that emerged in the literature was the impact of various technologies on teaching and learning. Miller and James [124] considered the potential learning benefits of adding animation techniques to previously static PowerPoint slides across two semesters of ASTRO 101 courses. To determine the effects of the animations, the authors administered analogous multiple-choice exams and pre- and postdiagnostic tests in course sections that used the animated slides versus sections that continued with static PowerPoint slides. Additionally, the instructors provided students with a Likert-style survey to “determine if there was any difference in perceived usefulness between the static and animated presentations” [124], p. 8]. Overall, students reported that they perceived the animated slides as more effective, despite learning comprehension measured through in-class exams remaining consistent between the courses that received the animated slides and those that received the static slides. Students in the animated slide courses, however, performed better on the diagnostic post-test, potentially suggesting the animations had some effect on retention of material. Duncan *et al.* [125] used a combination of observational techniques, surveys, and interviews to assess the potentially detrimental effects of a different type of technology, cell phone use, on student learning. Their mixed methods approach employed across

eight introductory astronomy and geology courses indicated that frequent in-class cell phone use was negatively correlated with final course grade, resulting in a drop of 0.366 ± 0.08 on a 4-point scale, where a 4.0 is an A. Their findings suggest that students were unable to both listen attentively and use their phones in class.

Yu and colleagues conducted three studies highlighting the effectiveness of visualizations (including different types) within planetarium-based instruction. Teaching the concept of seasons, Yu *et al.* [126] compared students' learning with the use of immersive planetarium-based visualizations ($n = 336$) to either flat (i.e., visualizations projected on a flat screen $n = 160$) or no visualizations ($n = 52$) among ASTRO 101 courses. They created a relevant test bank pulling from preinstructional student interviews and existing literature, including the ADT. Students experiencing the immersive visualizations scored significantly higher, with larger effect sizes, than did either other group; students who saw flat visualizations also scored higher than those who saw no visualizations, but the difference was not as stark. Similar results were found for instruction around moon systems within the solar system, with those students who experienced immersive visualizations scoring higher than the other two groups [127]. Immersive versus flat visualizations on the scale of the solar system were the focus of Yu *et al.* [128], and findings again supported the greater efficacy of immersive visualizations for learning. Fajriani and Masturi [129] used a pre- and post-test design to investigate astronomy students' learning about the celestial coordinate system using the Sky Map software (available in both web-based and app modes) and found that students' understanding did increase significantly over time.

Duncan and Arthurs [130] and Flohic [131] assessed the impact of new, embedded curriculum and overall course design on introductory astronomy students' attitudes toward science. In the Duncan and Arthurs [130] study, the research team compared two introductory astronomy courses, one taught traditionally and another with an embedded curriculum directly emphasizing the nature of science. The nature of science curriculum "gave students practice at evaluating examples of valid science and pseudoscience found on the internet; it also provided students opportunities to discuss what they [thought] about learning science" [130] (p. 1). Students in both courses were provided with an assessment called the Epistemological Beliefs Assessment for Physical Science (EBAPS; [132]) as well as a short-answer question on their final exam about the course's potential impact on their thinking about science. A subset of students from both classes was selected to participate in interviews that consisted of questions focused on the nature of science. Overall, students in the course with the embedded curriculum reported increased confidence in their ability to determine the credibility of scientific information from the internet. Compared to the traditional lecture course, students who

received the embedded curriculum were more inclined to perceive science as accessible to anyone, whereas those in the traditional course reported science was reserved for people with natural talent. Flohic [131] assessed introductory astronomy students' attitudes in the context of two entirely different courses: a traditional ASTRO 101 course and a Science, Technology, and Society (STS) seminar with more focused discussion, reading, and writing components. Most importantly, this study's results indicated that students in the ASTRO 101 course lost confidence in the general public's ability to understand science. Both studies lend credence to the idea that ASTRO 101 courses taught traditionally may require supplemental, targeted instruction aimed at improving students' perceptions of science as an accessible discipline, rather than focusing on content comprehension exclusively.

Additional efforts focused on instructional strategies or curricular products that were thematically different from the previous studies described so far. Garland and Ratay [133] conducted a pilot analysis of an introductory astronomy course that emphasized reading and writing assignments (including article and magazine responses) in favor of traditional quantitative-based problem sets. Using a subset of relevant ADT 2.0 items [46] in combination with their own test questions, the authors found that student assessment scores improved overall between the pre- and post-tests, although normalized gains were not calculated due to the small size of the course. Furthermore, students reported increased confidence with respect to their assessment answer choices between the pre- and post-tests, increasing from 16% (pretest) to 80% (post-test) of students reporting they were either confident or very confident with their answer choices. Similar increases were not reported, however, when students were asked their opinions regarding their ability to do science, suggesting that a single course (or even two courses) may not be enough time to significantly change students' perceptions of their scientific abilities.

During the large-scale validation study for the LSCI [64], Bardar and Brecher [134] noticed that one course section had particularly large gains. This course was using Project LITE, a series of six "homelabs" or activities designed to help students use inexpensive materials and their computers to investigate properties of light and spectra. Comparing this course ($n = 38$) with both lecture-based ($n = 1009$) and active learning via other strategies ($n = 683$) sections, students in the Project LITE course had mean post-test scores as high or higher than both other groups on all five subtopics. The authors concluded that the Project LITE materials were as or more effective than other evidence-based active learning strategies such as LTs and PI. Meyer *et al.* [135] introduced their Lunar Phases Project in ASTRO 101 courses to support students' mental model building. Students first documented their mental models on the cause of the Moon phases, then collected direct

observational data and revisited their models over time. Near the end of the project, they created visual models, along with a written explanation, that allowed for predictions that were then checked. Students completed the LPCI pre- and postproject. Over multiple semesters, significant increases with large effect sizes were seen from pre- to postproject.

3. Active learning in laboratory courses

One way in which the ASTRO 101 course varies across institutions is the inclusion of a laboratory component, and if present, its focus and goals. Some programs have no associated astronomy lab, for reasons that may range from philosophical and pedagogical approaches to credit requirements for both general education and major programs, as well as financial and logistical issues. For those where a lab is included, it may be focused on observational astronomy, such as through the use of small local or remote telescopes, or daytime labs that build knowledge and skills through simulations, data analysis, etc. There may also be broader primary or secondary goals, such as teaching critical thinking or problem-solving skills. As such, research on the efficacy or impact of labs is another area of AER. The studies that fall into this category focused on either a multi-instance treatment (incorporation of critical thinking skills or nighttime observing labs, individual versus group completion of inquiry activities) or single activity approaches (stereo displays for understanding lunar phases, simulation for understanding stellar parallax).

Allen and Kelly-Riley [136] described the process and initial results of the incorporation of specific critical thinking tasks into an existing astronomy lab. Faculty members involved with their university-wide critical thinking project scored both lab assignments and a random sample of student work from two semesters on a six-point rubric to assess the extent to which the lab assignments asked for and evoked, respectively, critical thinking skills. Early results were mixed, for example, showing that even when a lab explicitly required a demonstration of an aspect of critical thinking, students may have only implied it through their responses. Night-time observing labs were the focus of work by Jacobi *et al.* [137]. Activities were designed specifically for clear weather and to challenge known student difficulties, with the intent being for students to cycle through activity stations using a 16-inch telescope, a smaller telescope, or naked eye observations. A number of implementation inconsistencies and, of course, variable weather created challenges in analysis, but findings indicated an overall improvement in pre-to-post learning gains compared to the prior term's students on the same knowledge assessment. Sibbernson [138] looked at students' ability to engage with scientific inquiry when working individually compared to when working in cooperative lab groups. The mixed methods approach used the Views of Scientific Inquiry (VOSI; [139]) instrument

and lab activity data to assess any potential differences. The two treatment conditions fared similarly on the seven VOSI questions, with only one showing a different trend from pre-to-post administrations; an independent rater judged the individual condition having more detailed responses but chose "safer" or easier questions to investigate [138] (p. 49). Overall mean scores improved significantly from pre to post for the group condition but not for individuals.

Two different lab activities were created and assessed for their ability to help students better understand topics that require spatial thinking skills. Cid and Lopez [140] compared two lab activities for teaching lunar phases that were identical except for the use of a stereo visualization display in half of the sections. Assessed via the LPCI, the authors found that although students' scores were significantly higher at post-test than at pretest, there were no significant differences between the two groups, suggesting that stereo display may not offer any advantage over other visualization techniques. Ladd *et al.* [141] created a hybrid lab that included both hands-on and computer simulation (via the World Wide Telescope) techniques to teach parallax, with particular attention to its role as an astronomical distance measure (as opposed to only or primarily focusing on it as part of the definition of a parsec). Student understanding was measured by both the S3CI and a post-lab quiz. Results showed that although students had a good understanding of how parallax works on terrestrial scales (i.e., familiar settings and everyday phenomena), they still struggled after the lab activity to recognize the primacy of using it to measure astronomical distances.

4. Active learning in online courses

The last 20 years have seen dramatic changes in online learning as specific technological tools, Internet access, and student needs have evolved. As a result, there have been a number of studies that investigate varying online tools or course formats. Slater and Jones [142] applied the principles of active learning (or "learner-centered education") to an online course, creating interactive hypermedia tools that required students to first make their initial knowledge explicit and then to later compare their learning to their initial ideas. This online course was compared to a similarly designed in-person course. Students in the online course were able to successfully learn the content, but achievement by students in the in-person course was significantly higher.

Miller and Redman [143] created a series of video demonstrations for use in online courses and compared the performance of students in the online course with students in an in-person course that had the same demonstrations provided live. A number of exam questions relating to the demonstrations were given to both courses, and the ADT was used as an overall pre- and postassessment. Online course students who watched the videos (which were optional) performed significantly better on the associated exam questions than those who did not watch,

with no difference between either subset of online students compared to the in-person students. The in-person group fared significantly better than the online group on exam questions that did not relate to the demonstrations. In-person students also performed significantly higher on the ADT at posttest than did the online students, despite pretest scores being statistically indistinguishable. A related study investigated the extent to which the use of these video demonstrations related to students' perceptions of "instructor presence" [144]. As measured through both general and astronomy-specific surveys, students rated the videos effective at both creating instructor presence and at explaining astronomical concepts. There was also a positive correlation between perception of instructor presence and attitude toward the course; a lower but still positive correlation between perception of instructor presence and attitudes toward astronomy more generally; and a positive correlation between perceived helpfulness and perceived enjoyment of the videos. There was no significant correlation between the perceived helpfulness of the videos and performance on exam questions.

The use of Facebook as a place to host discussions in online courses served to elicit a greater number of initial posts and responses for an online course than a similar course that hosted discussions via the built-in learning management system (LMS; [145]). Students in the Facebook group tended to also post earlier and were more likely to post beyond the minimum required than their LMS counterparts. Barringer *et al.* [146] reported mixed results of an effort to improve student learning through "gamifying" an online astronomy course, which incorporated nonplayer characters and minigames in teaching the content. Using results from a subset of six TOAST questions most closely related to minigame content, two of the areas showed improvement, whereas four indicated that students continued to have difficulties (i.e., similar to a non-gamified approach). Horodyskyj *et al.* [147] assessed results from four semesters of a "habitable worlds" online course, outfitted with adaptive technology. Overall course grades were generally higher than most general education courses at the institution (50% A); successful and efficient completion of the course project was correlated with higher grades; and post-test attitudes on a relevant survey were generally higher than a benchmark provided by the survey developer. The authors also noted differences among various demographic groups that showed the course's effectiveness was not universal.

An online general education course in astrophysics was created by Trout *et al.* [148], with lectures created to go along with the OpenStax astronomy textbook [149]. A test created from textbook questions showed significant increases in knowledge over the semester. About half of the enrolled students also responded to a voluntary survey about the course. Although many felt the online course was effective and that they were likely to take additional online

courses, it is unclear whether this is representative of the full enrollment. Simon and colleagues [14] created a conceptual framework to develop and implement online activities that use large datasets from active citizen science projects intended to facilitate increased data literacy. They report on pilot testing at nine different universities with an assessment instrument given at pretest (10 items) and posttest (11 items, 7 of which overlapped with the pretest). Factor analysis demonstrated three factors in the instrument: citizen science, data literacy self-efficacy, and science engagement. Using matched pre- and postadministrations from before and after the "Planet Hunters" citizen science activity, Simon *et al.* found that students reported significant improvements in their knowledge; this was true for both the overall instrument and across the three factors. Interviews with instructors indicated a positive implementation experience, with many feeling that the incorporation of the activity in their course supported their learning goals and was a successful experience for the students.

E. Authentic research experiences

Providing undergraduates with opportunities to conduct authentic research has been deemed by the American Association of the Advancement of Science [150] and the National Academies [151] as one of the most essential and beneficial opportunities an undergraduate can have. Furthermore, some level of research experience is typically expected of undergraduates applying to graduate and professional schools, which is a common pre-career choice for many undergraduate science majors. AER has attempted to better understand the learning that occurs and the value of such experiences for students.

Course-based Undergraduate Research Experiences (CUREs) make research accessible to a broader population of learners, as research is conducted as part of a larger course, rather than through individual mentorship like traditional Research Experiences for Undergraduates (REUs). Studies on the successful implementation of CUREs in a variety of classroom contexts demonstrate a myriad of benefits for students and universities alike, including increased program retention, improved student data analysis skills, and increased scientific literacy [see, e.g., [152,153], and references therein]. Laboratory-based sciences like biology and chemistry have led the way in terms of CURE development and implementation, while the fields of astronomy and astrophysics have more recently followed suit.

Although not formally labeled a CURE, Lacy [154] discussed in detail the introduction of research-based mini projects as part of an introductory astronomy laboratory curriculum for non-science majors. Participating students used a web-based, robotic observatory named Undergraduate Research Studies in Astronomy (or URSA) to learn valuable observational astronomy skills such as target selection,

submitting observation requests, and conducting image analysis to estimate, for example, “the distance to a globular cluster by measuring its angular size” (p. 134). A preliminary analysis of Likert-style survey questions administered to students, along with the review of select student remarks and teaching assistant student evaluation scores, suggested that the URSA telescope was user-friendly and students were generally satisfied with the observational astronomy-based labs. Students also indicated increased satisfaction with the course and their teaching assistants when compared to previous semesters when the URSA labs were not included [154]. The implementation of robotic telescope and observational astronomy laboratory interventions in physics and astronomy classrooms laid the foundation for more extensive education research efforts that investigated the impact of these interventions on a variety of student affective outcomes (e.g., motivation, attitudes and beliefs about science, and student understanding of the nature of science; see [25] and references therein).

Wooten *et al.* [153] were the first to formally assess the impacts of course-based research experiences on undergraduate astronomy students in particular (rather than high school students or pre- or in-service teachers). They used a mixed methods approach to examine the effects of CUREs and Quasi CUREs (QCUREs) across a variety of student affective outcomes (e.g., motivation and attitudes toward doing science), with an emphasis on their perceived science process skills. QCUREs differ from traditional CUREs in that they do not include components of original scientific discovery or relevance to the field more broadly. In this study, students enrolled in introductory astronomy courses across four institutions of higher education participated in either semester-long research projects or shorter projects that “took place over two to four three-hour laboratories” [153] (p. 3). Ultimately, student participation in both CUREs and QCUREs led to increased confidence in students’ perceived science process skills, along with improvements in students’ motivation to learn science as well as engage in the process of doing science. CUREs outperformed QCUREs, however, when specifically considering the role of conducting analysis as part of the scientific process and the importance of collaboration in scientific discovery [153].

Hewitt *et al.* [152] built upon these efforts by developing and assessing one of the first online CUREs for astronomy majors. This CURE, referred to as the Exoplanet Research Experience (ERE), was offered as an elective course as part of a new, fully online, undergraduate degree program in Astronomical and Planetary Sciences. Students in the course worked collaboratively to update the orbital parameters of large, transiting exoplanets that could be observed using robotic, ground-based telescopes. A mixed-methods analysis including Likert-style surveys and student interviews indicated that participation in the course led to improved research self-efficacy, science identity, sense of belonging, and exoplanet content knowledge [152].

F. Studies with faculty and planetarium professionals

Although the primary focus in AER has been on students (e.g., their learning, motivation) or the class design or environment (e.g., use of active learning), select studies focused on the perspectives of faculty or instructors and planetarium professionals. Studies of faculty, in particular, are important in that there is no uniform instructional method used to teach astronomy, nor is there systematic training for astronomy faculty. In terms of training and expertise, planetarium professionals are likely more heterogeneous than astronomy faculty. By better understanding both of these populations, we can gain insights into how to support their teaching and professional growth, analogous to how we support student learning.

Fraknoi [155] conducted a survey of ASTRO 101-type instructors at community colleges (70% of the 400 responses) and teaching-intensive institutions (30%), with the purpose of better describing these faculty. Respondents were categorized by highest degree earned, field of said degree, institutional facilities available (e.g., laboratory or planetarium), class size, and available budget. The author provided a narrative description of “the life of a community college instructor,” although it is unclear whether the information here was culled from the survey or his personal knowledge of the community college environment. Two open-ended questions provided opportunities for respondents to describe “their proudest accomplishment” and the supports they wished they had for teaching astronomy.

Three studies looked in-depth at instruction within a single classroom. Lippert and Partridge [156] worked together in a student-instructor partnership, respectively, to investigate the extent to which enrolled students found the instructor’s goals of an ASTRO 101 course interesting and whether they were effectively met by the end of the semester. Although students felt many of the goals were met, some key goals (notably increased student confidence, familiarity with the night sky) were not sufficiently addressed. Brogt [157] described a study in which he participated as both a teaching assistant and mentor to a faculty member who engaged in professional development for changing his ASTRO 101 course from a traditional, teacher-centered approach to one that was more learner-centered. Their approach included having the instructor engage with the professional development as a “scientific experiment” and Brogt collected data via interviews, classroom observations, and research notes to document the instructor’s motivations for change, expectations for the course, reflection on the expectations post-instruction, and overall reflection on the process. Bailey and Nagamine [158] evaluated a similar change in an ASTRO 101 course (i.e., from more teacher- to more learner-centered) through the lens of Dole and Sinatra’s Cognitive Reconstruction of Knowledge Model [159] to understand the motivations, barriers, and supports, and perceived satisfaction of the change. They also looked at student learning as measured

by the SPCI, finding that students in the revised course scored significantly higher on the posttest than did those who were in the traditionally designed course.

A number of astronomy courses may take advantage of opportunities afforded by the existence of a campus planetarium. Everding and Keller [160,161] conducted two mixed methods studies to investigate planetarium use in undergraduate astronomy coursework. The earlier study [160] focused on understanding the nature of the planetarium (size, projector type, courses, etc.) and its use (e.g., content, interactive strategies used) through surveys and interviews of planetarium representatives. Many planetariums' primary use was with local K-12 learners rather than undergraduates, and undergraduate use was not limited to astronomy courses. The later study [161] focused on how and why faculty at a single institution used (or did not use) a planetarium environment in their instructional design, as well as the costs, benefits, and norms of said use, for undergraduate astronomy courses through surveys, interviews, and class observations. Findings indicated that use of the planetarium may depend upon the instructor's knowledge of the facility's capabilities and that it is often used with more affective goals (as opposed to cognitive or learning goals) in mind.

Although some astronomy courses make use of planetaria for part or all of their instruction (see, e.g., [160]), these environments are also popular with K-12 students and the public. Planetarium professionals, similarly, may come with a significant astronomy background or may have been trained in other areas. Plummer and Small [162] interviewed planetarium professionals about the goals for their audiences (both general audiences and shows aimed primarily at children; the latter was addressed more fully in a companion publication [163]) as well as their design choices. Using six lenses elucidated in two National Research Council publications [164,165], the researchers found that the participants' top two goals for general audiences were to provide educational content and to create interest or engagement with the content. All of the participants provided responses that aligned with at least two of the six lenses, with the knowledge-oriented and motivation-oriented lenses being the most commonly represented, in line with the stated goals.

Schultz and Slater [166] attempted a first-pass look at planetarium professionals through an analysis of background information self-reported by 61 individuals on an open Facebook group. They described the group in terms of education (e.g., 93% bachelor's, 38% masters, and 10% doctoral degrees) and prior experience (e.g., up to a third each with museum or informal education experience, a teaching certificate or degree, or performance experience; these categories may not be exclusive to one another). These authors also analyzed audio recordings of planetarium shows to determine to what extent the presenters used a formative assessment cycle (in which the presenter elicits a

response from the audience and uses that response to determine support learning in subsequent actions), followed by interviews to understand the barriers to using active learning strategies in the planetarium [167]. More than one-third of the recordings were found to contain no formative assessment cycles and another third included only one to two instances. Interviews documented engagement or entertainment as a primary goal, similar to the findings of Plummer and Small [162].

G. Studies with adult learners more broadly

Astronomy appears to be one of the few science areas where amateurs can have a large involvement in public outreach and contribute to research (e.g., via the *Journal of the American Association of Variable Star Observers*). Astronomy is also popular for adults enrolling in Massive Open Online Courses or engaging with citizen science efforts. Better understanding aspects of these groups—who they are, why they engage in the way they do, and what they get out of it—has been the focus of some AER studies.

Berendsen [168] set out to determine the astronomy content knowledge of this population, using a survey that included ADT questions and demographics, and advertised to officers of astronomy clubs across the United States. With more than 1100 complete and valid responses, the mean ADT score was 76% with a range of 66% (no formal astronomy education) to 92% (astronomy, astrophysics, or physics degree). Results were further broken down by years of membership in an astronomy club, outreach frequency, and combinations of the three primary groupings. From two surveys of amateur astronomers, Yocco *et al.* [169] found that both individual (level of formal or informal training and age) and club-related (club membership, length of membership, club service) factors were related to an individual's likelihood of engaging in education and public outreach.

Outreach efforts have also been of interest to a number of researchers. Price *et al.* [170] report on an evaluation of the *Slacker Astronomy* podcast, using interviews followed by surveys to determine listeners' attitudes toward the podcast and astronomy knowledge relating to specific shows. Three studies looked at citizen scientists—their motivations for participating in a Zooniverse project or their conceptual understanding of astronomy. A survey of just under 11 000 Galaxy Zoo participants about the presence of 12 different categories of motivation for engaging with the platform indicated that the most frequently chosen primary motivation for both men and women was to contribute to science (though an earlier question allowed participants to select multiple motivations; [171]). Prather *et al.* [172] administered content knowledge surveys to both Galaxy Zoo and Moon Zoo users, finding that the highest-engaging users also had the highest survey scores (although the authors caution that they make no claims on the extent to which participation helps users learn this content). Nearly 1500

Galaxy Zoo users responded to a content knowledge survey by Raddick *et al.* [173], with the results then analyzed to look for relationships between knowledge scores and three different measures of usage (number of galaxies classified or two time-based measures of involvement). Effects for all were small, suggesting that involvement in Galaxy Zoo classification did not influence participants' conceptual understanding.

Related to—but different from—both outreach efforts described in this section and online courses (Sec. IV D 4), are efforts around Massive Open Online Courses, or MOOCs, in astronomy. Chen *et al.* [174] used a survival analysis approach to understand how prior knowledge influenced MOOC-takers' persistence in the course. In general, students with lower pretest scores (i.e., students with more alternative conceptions) were more likely to drop out of the course earlier, though this effect was reduced over time. The authors identified a number of variables that positively influenced a student's completion chances as well as others that had no impact.

Impey and Formanek [175] report on how astronomy MOOC enrollment changed from before to during the early months of the COVID-19 pandemic. Using course enrollment data from two MOOC providers and a motivation survey administered in courses from one of them, the authors found that enrollments increased dramatically in a short period of time, that new learners tended to be younger than past enrollees, and that many were full-time undergraduate students. Particularly large numbers of students came from India and developing countries in Asia. Impey *et al.* [176] compared general knowledge and attitudes, and beliefs about science of astronomy MOOC learners to those of undergraduates nonscience majors, and to science experts. MOOC learners tended to score slightly higher than undergraduates on a science literacy test and gave longer answers on open-ended questions about science; additionally, the length of responses was correlated with science literacy score. There were also differences in some of the specific topics (e.g., MOOC learners performed better on questions about DNA and software) and in general attitudes and faith-based beliefs across the three groups.

V. DISCUSSION

A. The current state of AER

We identified 135 articles for inclusion in this review and have categorized them into seven groups (see Table I for the number of articles in each category). In addition to the research articles, we identified many “how-to” and curricular overview articles that, while valuable to the astronomy education literature more broadly, do not meet the inclusion criteria specified in Sec. III A. Findings discussed in a recent report by Salimpour *et al.* [177] note that curriculum and/or program descriptions or reports

make up 43% of the astronomy education papers published between 1898 and 2022. Empirical research is the second-most popular category at 37%, although the authors do not define or provide specific examples within their categories. Additionally, Salimpour *et al.* noted a decline in the number of astronomy education articles published since 2017, with a nearly 50% decrease between 2018 and 2019 in particular. The publication rate has continued to decline and then plateaued post-COVID. The authors attribute this decline to several factors, including the pandemic's interference with the ability to collect data with teachers and students, and the potential for publication in more “obscure” journals that had not yet made it into the database they used. Regardless of the reasons, this decline is concerning and emphasizes the importance of continuing to publish quality AER articles.

B. Moving beyond Bailey and Slater's review

While the goal of this study was not to recreate the Bailey and Slater [3] review, we use it as a starting point for initial comparisons to better understand how the field has evolved over the last two decades. Bailey and Slater [3] binned articles into three primary categories in their review: (a) reviews and status papers; (b) research on student (and teacher) understanding of astronomy topics; and (c) research on various instructional methods. These three categories continued to be strongly influential, although—as would be expected—there have been shifts in the details.

1. Review papers

Researchers have continued to compile and synthesize research in reviews, with 12 identified, but in addition to broader status reviews such as this one, sufficient works exist in some topic areas that more focused reviews have been written (e.g., gravity [19,20]; spatial thinking [24]). Additionally, authors have provided retrospectives about particular AER-publishing venues [16,17], something that had not been possible in the past. These review papers can be particularly helpful for new astronomy education researchers, providing an entry point into both the broad field and specific focal areas. AER review papers thus far (the present study included) have been narrative in nature, focusing on either a broad state of the field or a particular topic. What we have not yet seen is a more systematic review, which is defined by Alexander [2] as a process and resulting product that “entails an orchestrated search of literature intended to address a *particular question or issue of importance to the field* ... and the conclusions and implications that arise from this process are, thus, well-grounded in the examined literature” (p. 8). Relatedly, AER has not yet attempted any quantitative meta-analyses that would compare findings across studies in an effort to draw larger conclusions.

2. Student understanding

Research on student understanding has remained popular, with 25 articles included in this review. Two shifts in this area over the last 20 years are of particular note. First, the topics of focus have broadened to include more modern and complex topics as befitting astronomy courses for undergraduate students; such studies are not as focused on content that tends to be addressed in the K-12 curriculum as they had been in the past. This includes a number of works on, for example, cosmology [32–35,38] or the nature of time [23]. This claim is not based upon the exclusion of K-12 research in the present review; rather, it is that AER at the tertiary level has focused on different content in more recent years than in the past. Second, the analysis of student preinstructional understanding has moved away from a “target misconceptions for correcting” approach to researchers (and instructors) recognizing that students’ ideas can serve as a resource for learning rather than exclusively as a barrier (cf., [178]). This follows trends in science education research more broadly, which no longer sees a cataloging of student difficulties as a primary need [179]. Bailey and Slater [3] noted that “a second area of [needed] research ... is the underlying cause of student difficulties in astronomy” (p. 37). Future studies in this area should not simply identify alternative conceptions, but look at the sources and mechanisms of such difficulties—for example, to what extent a student’s learning about astronomical motions may be impacted by their more general spatial thinking skills [24]. We offer a reminder here that research on teacher (preservice or inservice) understanding, which was prominent in the reviews by both Bailey and Slater [3] and Lelliott and Rollnick [8], has been deliberately excluded from the present analysis and will be covered in a future paper.

3. Instructional strategies

Finally, our largest category continued to be research on the efficacy of various instructional strategies, with 45 articles identified. With a large adoption rate in ASTRO 101 courses, LTs, RTs, and TPS/PI questions have dominated this space. Many of the papers included here focused on the effectiveness of implementing one or more of these strategies in different contexts; this is particularly apparent in papers published between 2004–2014. All three of these are seen as active learning strategies [92] and, despite remaining popular, have yielded mixed results in the research literature. This may be influenced, for example, by the specific classroom implementation details, what assessment is being used (e.g., the ADT versus a customized assessment), and when postinstructional understanding is measured (e.g., immediately after a single LT versus at the end of the course). Additional active learning strategies have received some attention, as have individual laboratory activities (e.g., [140]) and lab course redesigns (e.g., incorporating inquiry, [138]). Over time, we see more

AER branching into technological approaches, such as strategies involving simulations (e.g., [141]) or incorporating active learning strategies into fully online courses (e.g., [14]). Online astronomy courses are not new, but advances in technology in general have opened this area to more students and, subsequently, more researchers. As noted in the Bailey and Slater [3] review, more research is needed on “...the practice–theory connection that makes use of the theory-driven research in real classrooms” (p. 37). In the present review, we identified many studies on instructional effectiveness. A few of them attempt to apply theory to the undergraduate classroom. As one example, Simon *et al.* [14] developed a curricular model based on the theory of social semiotics, developing new representations of astronomical content to bridge everyday understandings that students bring to the classroom with discipline-specific presentation of content. However, most of the AER studies reviewed here—along with similar studies in other content areas—often still lack the more meaningful connections to the theoretical framework behind why specific instructional strategies or curricular interventions actually work. Lombardi *et al.* [92] noted that astronomy is like most other discipline-based education research fields in that the definition of active learning has largely been treated as “anything but lecture”; this may also be a contributor to mixed results in this area. By applying the Construction of Understanding Ecosystem: Active Learning framework developed by Lombardi *et al.* [92]—or another framework—to design active learning environments, we may be able to better explain *why* the strategies work, under what *conditions* or in what *contexts* they are most effective, and *who* will directly benefit.

Bailey and Slater [3] also suggested “... the need for studies on student variables and their impact on student understanding” (p. 37). Education research is often conducted with convenience samples (i.e., our own students), and this may not be representative of the global community [180]. Thus, when we look at student understanding or instructional effectiveness, we may have an incomplete picture of both successes and needs. Expanding the reach of our research may involve, for example, partnering with more researchers in international settings or looking at ways in which different groups of learners engage with astronomy content through aspects such as cultural heritage or personal relevance.

4. New categories

We also identified four new categories of AER that were not identified by Bailey and Slater [3]. In some cases, this is a function of how we grouped the studies (for example, the 2003 review includes two assessment papers within the student understanding category, whereas we separate them here). In others, the categories are completely new and distinct from past work. While it might be possible to maintain the original three categories from the 2003 review,

our intent was to identify the recent and current areas of focus in the field, regardless of their alignment (or not) with past reviews.

a. Assessment tools. Our first new area, assessment of learning and other outcomes, was our second largest category, with 31 articles included in the present review. The original 2003 review included a subsection on the development of the ADT [46], and the LPCI [60] was described in the student's understanding of the Moon subcategory. Since then, researchers have developed and tested a number of additional concept inventories (e.g., SPCI [61,62]; PFCI [69]), as well as instruments to measure affective outcomes such as self-efficacy [81,82]. A motivation for the development of concept inventories, in particular, has been to have an analog to the FCI [58,59] used in physics that can allow for comparisons across classes and instructional strategies. Such use, however, does not appear to have played out writ large across the field. Given the nature of ASTRO 101 as (typically) a stand-alone, general education course—rather than the foundation of other programs of study, which is true for most introductory physics courses—it may not be realistic to expect any one assessment to provide the kinds of insights that the FCI has. That said, one of the challenges in doing systematic reviews and meta-analyses is in how outcomes are assessed, and common instruments with a targeted approach could be one way to further these areas of AER.

b. Authentic research experiences. The three other new categories are notably smaller than the others described so far, but expand the field of AER in interesting and meaningful ways. One of the newest, and therefore smallest, areas of AER is in the use and analysis of incorporating authentic research experiences for students into astronomy courses, such as through CUREs [152,153]. This is an example of an instructional strategy that is well-aligned with the framework developed by Lombardi *et al.* [92] (p. 18), allowing students to engage in direct experiences with science, including domain practices, data, and models, with both peers and the teacher, rather than receiving such information second-hand with the teacher acting as a filter. We describe studies on CUREs and undergraduate research experiences in Sec. IV E, some of which highlight the use of remote or robotic telescopes that allow students to request astronomical observations that they can then analyze. Fitzgerald *et al.* [25] reviewed this work as part of the 2017 Robotic Telescopes, Student Research, and Education Conference Proceedings, but education research was only a small component of the review. More often, the papers associated with this topic have described the technology and its educational uses [181,182]. Detailed research on outcomes of robotic telescope use, specifically, whether cognitive or affective, has only recently begun to appear (e.g., [183], which was published during the writing

of this paper but fell outside of our inclusion time period). Separately, expansion on the potential benefits of CUREs, with particular attention to student demographics and both cognitive and affective outcomes, would also improve this area of research.

c. Faculty, planetarium professionals, and adult learners. Studies with faculty and planetarium professionals, and with adults more broadly (excepting K-12 teachers), our final two new categories, are other areas of growth. Although we often assume little difference between undergraduate students and adults, the reality is that there are significant changes in the brains—reasoning, functioning, etc.—and experiences of young adults during the nominal college period [184], and differences may be hidden by our assumptions. Additionally, focusing on the role of instructors, whether faculty [157,158,160] or planetarium professionals [160], may provide greater insights into the motivations for and mechanics of instructional change.

VI. CONCLUSIONS

While progress has certainly been made in the field of AER over the last two decades, we should note that AER is sometimes viewed as lagging behind our peer communities, such as physics education research. This is, in all likelihood, an issue of scale: Although ASTRO 101 courses are incredibly popular and widespread in the United States [185], there are many fewer astronomy education researchers than there are physics education researchers. Additionally, the nature and number of astronomy courses vary considerably from those in physics, and so parallel research pathways may be neither feasible nor desirable. Instead, we should focus on the unique educational opportunities the field of astronomy affords. This might include, for example, introducing space-based data—often in the form of large datasets—to the classroom in an accessible way or leveraging the widespread popularity of astronomy as a general education science course. While the context of ASTRO 101 is often a mainstay of AER, studies that specifically engage with the goals of general education programs or broader science literacy research (such as that based in K-12 and informal science education) are worthwhile areas of future investigation. Better understanding these areas of research and how to leverage students' and the public's interest might help astronomers make the field more accessible to a greater number of people on the level that they desire.

ACKNOWLEDGMENTS

Publication of this article was funded in part by the Temple University Libraries Open Access Publishing Fund.

J. M. B. and M. N. S. contributed equally to this work.

DATA AVAILABILITY

The data are not publicly available. The data are available from the authors upon reasonable request.

APPENDIX

TABLE II. All of the summarized papers included in this review along with their reference number (R No.), category, and subcategory (if applicable).

Reference with R No.	Category	Subcategory
Agan [29]	Studies on Student Understanding	Understanding of Astronomical Content
Alexander [99]	Efficacy of Active Learning Approaches	The Big 3
Allen and Kelly-Riley [136]	Efficacy of Active Learning Approaches	Active Learning in Laboratory Courses
Bailey [61]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Bailey <i>et al.</i> [30]	Studies on Student Understanding	Understanding of Astronomical Content
Bailey and Nagamine [158]	Studies with Faculty and Planetarium Professionals	
Bailey <i>et al.</i> [32]	Studies on Student Understanding	Understanding of Astronomical Content
Bailey <i>et al.</i> [62]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Bailey and Lombardi [9]	Review papers	
Bailey <i>et al.</i> [86]	Assessment of Learning and Other Outcomes	Assessing Affective Outcomes in Astronomy
Bailey <i>et al.</i> [44]	Studies on Student Understanding	Understanding of Astronomical Content
Bardar [80]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Bardar <i>et al.</i> [64]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Bardar and Brecher [134]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Barringer <i>et al.</i> [146]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Bekaert <i>et al.</i> [74]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Berendsen [168]	Studies with Adult Learners More Broadly	
Bergstrom <i>et al.</i> [91]	Assessment of Learning and Other Outcomes	Assessing Affective Outcomes in Astronomy
Berryhill and Slater [87]	Assessment of Learning and Other Outcomes	Assessing Affective Outcomes in Astronomy
Blue [55]	Studies on Student Understanding	Understanding of Broader Aspects of Science
Bretones and Megid Neto [15]	Review papers	
Bretones <i>et al.</i> [17]	Review papers	
Brock <i>et al.</i> [23]	Review papers	
Brogt [157]	Studies with Faculty and Planetarium Professionals	
Brogt <i>et al.</i> [79]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Buck Bracey [41]	Studies on Student Understanding	Understanding of Astronomical Content
Buck [40]	Studies on Student Understanding	Understanding of Astronomical Content
Buxner <i>et al.</i> [88]	Assessment of Learning and Other Outcomes	Assessing Affective Outcomes in Astronomy
Chen <i>et al.</i> [174]	Studies with Adult Learners More Broadly	
Cid and Lopez [140]	Efficacy of Active Learning Approaches	Active Learning in Laboratory Courses
Coble <i>et al.</i> [33]	Studies on Student Understanding	Understanding of Astronomical Content
Coble <i>et al.</i> [34]	Studies on Student Understanding	Understanding of Astronomical Content

(Table continued)

TABLE II. (Continued)

Reference with R No.	Category	Subcategory
Coble <i>et al.</i> [37]	Studies on Student Understanding	Understanding of Astronomical Content
Cole <i>et al.</i> [24]	Review papers	
Conlon <i>et al.</i> [36]	Studies on Student Understanding	Understanding of Astronomical Content
Crider [113]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Duncan and Arthurs [130]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Duncan <i>et al.</i> [125]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Eckenrode <i>et al.</i> [102]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Eriksson <i>et al.</i> [39]	Studies on Student Understanding	Understanding of Astronomical Content
Everding and Keller [160]	Studies with Faculty and Planetarium Professionals	
Everding and Keller [161]	Studies with Faculty and Planetarium Professionals	
Fajriani and Masturi [129]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Favia <i>et al.</i> [75]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Favia <i>et al.</i> [76]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Fitzgerald <i>et al.</i> [25]	Review papers	
Flohic [131]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Fraknoi [155]	Studies with Faculty and Planetarium Professionals	
Fraknoi [16]	Review papers	
Freed <i>et al.</i> [81]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Freed <i>et al.</i> [82]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Garland and Ratay [133]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Gross and Lopez [42]	Studies on Student Understanding	Understanding of Astronomical Content
Hewitt <i>et al.</i> [152]	Authentic Research Experiences	
Heyer <i>et al.</i> [50]	Studies on Student Understanding	Understanding of Astronomical Content
Hintz <i>et al.</i> [31]	Studies on Student Understanding	Understanding of Astronomical Content
Horodyskyj <i>et al.</i> [147]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Hudgins <i>et al.</i> [110]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Impey and Formanek [175]	Studies with Adult Learners More Broadly	
Impey <i>et al.</i> [176]	Studies with Adult Learners More Broadly	
Jacobi <i>et al.</i> [137]	Efficacy of Active Learning Approaches	Active Learning in Laboratory Courses
James <i>et al.</i> [115]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Kavanagh <i>et al.</i> [18]	Review papers	
Kavanagh and Sneider [19]	Review papers	
Kavanagh and Sneider [20]	Review papers	
Lacy [154]	Authentic Research Experiences	
Ladd [141]	Efficacy of Active Learning Approaches	Active Learning in Laboratory Courses
Len [114]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Lippert and Partridge [156]	Studies with Faculty and Planetarium Professionals	

(Table continued)

TABLE II. (*Continued*)

Reference with R No.	Category	Subcategory
LoPresto [78]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
LoPresto [104]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
LoPresto and Hubble-Zdanowski. [70]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
LoPresto and Murrell [101]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
LoPresto and Slater [108]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Madsen [21]	Review papers	
Meyer [135]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Miller [145]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Miller and James [124]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Miller and James [123]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Miller and Redman [143]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Miller and Redman [144]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Nottis <i>et al.</i> [68]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Plummer and Small [162]	Studies with Faculty and Planetarium Professionals	
Prather and Brissenden [116]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Prather <i>et al.</i> [97]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Prather <i>et al.</i> [94]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Prather <i>et al.</i> [103]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Prather <i>et al.</i> [172]	Studies with Adult Learners More Broadly	
Price <i>et al.</i> [170]	Studies with Adult Learners More Broadly	
Pundak [77]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Pundak <i>et al.</i> [53]	Studies on Student Understanding	Understanding of Astronomical Content
Raddick <i>et al.</i> [171]	Studies with Adult Learners More Broadly	
Raddick <i>et al.</i> [173]	Studies with Adult Learners More Broadly	
Richards and Kelly [85]	Assessment of Learning and Other Outcomes	N/A
Robertson and Finch [84]	Assessment of Learning and Other Outcomes	Assessments of Additional Variables
Rudolph [118]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Rudolph <i>et al.</i> [96]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Schlingman <i>et al.</i> [65]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Schultz and Slater [166]	Studies with Faculty and Planetarium Professionals	N/A
Schultz and Slater [167]	Studies with Faculty and Planetarium Professionals	N/A
Sibbersen [138]	Efficacy of Active Learning Approaches	Active Learning in Laboratory Courses
Simon <i>et al.</i> [43]	Studies on Student Understanding	Understanding of Astronomical Content

(*Table continued*)

TABLE II. (Continued)

Reference with R No.	Category	Subcategory
Simon <i>et al.</i> [69]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Simon <i>et al.</i> [14]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Simonelli and Pilachowski [49]	Studies on Student Understanding	Understanding of Astronomical Content
Slater [52]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Slater <i>et al.</i> [51]	Studies on Student Understanding	Understanding of Astronomical Content
Slater and Jones [142]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Slater and Tatge [22]	Review papers	
Small and Plummer [163]	Studies with Faculty and Planetarium Professionals	
Straits and Wilke [120]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Sugarman <i>et al.</i> [54]	Studies on Student Understanding	Understanding of Broader Aspects of Science
Trouille <i>et al.</i> [35]	Studies on Student Understanding	Understanding of Astronomical Content
Trout <i>et al.</i> [148]	Efficacy of Active Learning Approaches	Active Learning in Online Courses
Wallace and Bailey [63]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Wallace <i>et al.</i> [71]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Wallace <i>et al.</i> [72]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Wallace <i>et al.</i> [73]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Wallace <i>et al.</i> [38]	Studies on Student Understanding	Understanding of Astronomical Content
Wallace <i>et al.</i> [105]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Wallace <i>et al.</i> [89]	Assessment of Learning and Other Outcomes	Assessing Affective Outcomes in Astronomy
Wallace <i>et al.</i> [107]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Wallace <i>et al.</i> [66]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Wallace <i>et al.</i> [106]	Efficacy of Active Learning Approaches	“The Big 3”: Lecture-Tutorials, Ranking Tasks, and Think-Pair-Share
Williamson and Willoughby [28]	Studies on Student Understanding	Understanding of Astronomical Content
Williamson <i>et al.</i> [67]	Assessment of Learning and Other Outcomes	Instrument Development for Astronomical Content
Wooten <i>et al.</i> [83]	Assessment of Learning and Other Outcomes	Assessments of Additional Variables
Wooten <i>et al.</i> [153]	Authentic Research Experiences	
Yocco <i>et al.</i> [169]	Studies with Adult Learners More Broadly	
Yu <i>et al.</i> [27]	Studies on Student Understanding	Understanding of Astronomical Content
Yu <i>et al.</i> [126]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Yu <i>et al.</i> [127]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Yu <i>et al.</i> [128]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom
Zeilik and Morris [45]	Studies on Student Understanding	Understanding of Astronomical Content
Zeilik and Morris [122]	Efficacy of Active Learning Approaches	Additional Active Learning Interventions in the In-Person Classroom

- [1] R. F. Baumeister, Writing a literature review, in *The Portable Mentor: Expert Guide to a Successful Career in Psychology*, edited by M. J. Prinstein (Springer, New York, NY, 2013), p. 119.
- [2] P. A. Alexander, Methodological guidance paper: The art and science of quality systematic reviews, *Rev. Educ. Res.* **90**, 6 (2020).
- [3] J. M. Bailey and T. F. Slater, A review of astronomy education research, *Astron. Educ. Rev.* **2**, 20 (2003).
- [4] G. S. Hall and C. E. Browne, Children's ideas of fire, heat, frost and cold, *Pedagog. Seminary* **10**, 27 (1903).
- [5] G. W. Haupt, First grade concepts of the Moon, *Sci. Educ.* **32**, 258 (1948).
- [6] G. W. Haupt, First grade concepts of the Moon: Part II: By interview, *Sci. Educ.* **34**, 224 (1950).
- [7] C. A. Wall, A review of research related to astronomy education, *Sch. Sci. Math.* **73**, 653 (1973).
- [8] A. Lelliott and M. Rollnick, Big ideas: A review of astronomy education research 1974–2008, *Int. J. Sci. Educ.* **32**, 1771 (2010).
- [9] J. M. Bailey and D. Lombardi, Blazing the trail for astronomy education research, *J. Astron. Earth Sci. Educ.* **2**, 77 (2015).
- [10] S. King, K. Davison, A. Chitiyo, and D. Apple, Evaluating article search and selection procedures in special education literature reviews, *Remedial Spec. Educ.* **41**, 3 (2020).
- [11] J. J. Birriel, Demonstrating absorption spectra using commercially available incandescent light bulbs, *Astron. Educ. Rev.* **7**, 147 (2008).
- [12] C. D. Impey, K. K. Hardegree-Ullman, A. Patikkal, A. Srinathan, C. L. Austin, N. K. Ganesan, and B. C. Guevenen, A new online astronomy resource for education and outreach, *Astron. Educ. Rev.* **12**, 010301 (2013).
- [13] C. S. Wallace, T. G. Chambers, E. E. Prather, and G. Brissenden, Using graphical and pictorial representations to teach introductory astronomy students about the detection of extrasolar planets via gravitational microlensing, *Am. J. Phys.* **84**, 335 (2016).
- [14] M. N. Simon, E. E. Prather, I. S. Rosenthal, M. Cassidy, J. K. L. Hammerman, and L. Trouille, A new curriculum development model for improving undergraduate students' data literacy and self-efficacy online astronomy classrooms, *Astron. Educ. J.* **2**, 1 (2022).
- [15] P. S. Bretones and J. Megid Neto, An analysis of papers on astronomy education in proceedings of IAU meetings from 1988 to 2006, *Astron. Educ. Rev.* **10**, 010102 (2011).
- [16] A. Fraknoi, A brief history of publishing papers on astronomy education research, *J. Astron. Earth Sci. Educ.* **1**, 37 (2014).
- [17] P. S. Bretones, L. C. Jafelice, and J. E. Horvath, Ten years of Latin-American Journal of Astronomy Education RELEA: Achievements and challenges for international astronomy education development, *J. Astron. Earth Sci. Educ.* **3**, 111 (2016).
- [18] C. Kavanagh, L. Agan, and C. I. Sneider, Learning about phases of the Moon and eclipses: A guide for teachers and curriculum developers, *Astron. Educ. Rev.* **4**, 19 (2005).
- [19] C. Kavanagh and C. I. Sneider, Learning about gravity I. Free fall: A guide for teachers and curriculum developers, *Astron. Educ. Rev.* **5**, 21 (2007).
- [20] C. Kavanagh and C. I. Sneider, Learning about gravity II. Trajectories and orbits: A guide for teachers and curriculum developers, *Astron. Educ. Rev.* **5**, 53 (2007).
- [21] A. Madsen, S. B. McKagan, and E. C. Sayre, Resource letter RBAI-1: Research-based assessment instruments in physics and astronomy, *Am. J. Phys.* **85**, 245 (2017).
- [22] T. F. Slater and C. B. Tatge, *Research on Teaching Astronomy in the Planetarium* (Springer International Publishing, Cham, 2017).
- [23] L. S. Brock, E. Prather, and C. Impey, Finding the time: Exploring a new perspective on students' perceptions of cosmological time and efforts to improve temporal frameworks in astronomy, *Phys. Rev. Phys. Educ. Res.* **14**, 010138 (2018).
- [24] M. Cole, C. Cohen, J. Wilhelm, and R. Lindell, Spatial thinking in astronomy education research, *Phys. Rev. Phys. Educ. Res.* **14**, 010139 (2018).
- [25] M. Fitzgerald, R. Cutts, S. Salimpour, and S. Slater, A brief overview of robotic telescopes, student research, and education research in the English literature, in *Robotic Telescopes, Student Research, and Education (RTSRE) Proceedings*, edited by M. T. Fitzgerald *et al.* (RTSRE, San Diego, CA, 2018).
- [26] T. G. Amin, C. L. Smith, and M. Wiser, Student conceptions and conceptual change, in *Handbook of Research on Science Education*, edited by N. G. Lederman and S. K. Abell (Routledge, New York, NY, 2014), p. 57.
- [27] K. C. Yu, K. Sahami, and G. Denn, Student ideas about Kepler's laws and planetary orbital motions, *Astron. Educ. Rev.* **9**, 010108 (2010).
- [28] K. E. Williamson and S. D. Willoughby, Student understanding of gravity in introductory college astronomy, *Astron. Educ. Rev.* **11**, 010105 (2012).
- [29] L. Agan, Stellar ideas: Exploring students' understanding of stars, *Astron. Educ. Rev.* **3**, 77 (2004).
- [30] J. M. Bailey, E. E. Prather, B. Johnson, and T. F. Slater, College students' preinstructional ideas about stars and star formation, *Astron. Educ. Rev.* **8**, 010110 (2009).
- [31] E. G. Hintz, M. L. Hintz, and M. J. Lawler, Prior knowledge base of constellations and bright stars among non-science majoring undergraduates and 14–15 year old students, *J. Astron. Earth Sci. Educ.* **2**, 115 (2015).
- [32] J. M. Bailey, K. Coble, G. L. Cochran, D. M. Larrieu, R. Sanchez, and L. R. Cominsky, A multi-institutional investigation of students' preinstructional ideas about cosmology, *Astron. Educ. Rev.* **11**, 010302 (2012).
- [33] K. Coble, C. T. Camarillo, L. E. Trouille, J. M. Bailey, G. L. Cochran, M. D. Nickerson, and L. R. Cominsky, Investigating student ideas about cosmology I: Distances and structure, *Astron. Educ. Rev.* **12**, 010102 (2013).
- [34] K. Coble, M. D. Nickerson, J. M. Bailey, L. E. Trouille, G. L. Cochran, C. T. Camarillo, and L. R. Cominsky, Investigating student ideas about cosmology II: Composition of the Universe, *Astron. Educ. Rev.* **12**, 010111 (2013).
- [35] L. E. Trouille, K. Coble, G. L. Cochran, J. M. Bailey, C. T. Camarillo, M. D. Nickerson, and L. R. Cominsky,

- Investigating student ideas about cosmology III: Big Bang theory, expansion, age, and history of the Universe, *Astron. Educ. Rev.* **12**, 010110 (2013).
- [36] M. Conlon, K. Coble, J. M. Bailey, and L. R. Cominsky, Investigating student ideas about the fate of the Universe, *Phys. Rev. Phys. Educ. Res.* **13**, 020128 (2017).
- [37] K. Coble, M. Conlon, and J. M. Bailey, Investigating undergraduate students' ideas about the curvature of the Universe, *Phys. Rev. Phys. Educ. Res.* **14**, 010144 (2018).
- [38] C. S. Wallace, E. E. Prather, and D. K. Duncan, A study of general education astronomy students' understandings of cosmology. Part IV. Common difficulties students experience with cosmology, *Astron. Educ. Rev.* **11**, 010104 (2012).
- [39] U. Eriksson, C. Linder, J. Airey, and A. Redfors, Who needs 3D when the Universe is flat?, *Sci. Educ.* **98**, 412 (2014).
- [40] Z. Buck, The effect of color choice on learner interpretation of a cosmology visualization, *Astron. Educ. Rev.* **12**, 010104 (2013).
- [41] Z. E. Buck Bracey, Personal universes: Revealing community college students' competences through their organization of the cosmos, *Cult. Stud. Sci. Educ.* **13**, 925 (2018).
- [42] N. A. Gross and R. E. Lopez, Advanced undergraduate and early graduate physics students' misconception about solar wind flow: Evidence of students' difficulties in distinguishing paradigms, *Astron. Educ. Rev.* **8**, 010102 (2009).
- [43] M. N. Simon, S. Buxner, and C. Impey, A survey and analysis of college students' understanding of planet formation before instruction, *Astrobiology* **18**, 1594 (2018).
- [44] J. M. Bailey, S. D. Willoughby, and B. A. Samie, An exploration of students' understanding of space exploration, *Astron. Educ. J.* **3**, 061aer-1 (2023).
- [45] M. Zeilik and V. J. Morris, An examination of misconceptions in an astronomy course for science, mathematics, and engineering majors, *Astron. Educ. Rev.* **2**, 101 (2003).
- [46] B. Hufnagel, Development of the astronomy diagnostic test, *Astron. Educ. Rev.* **1**, 47 (2002).
- [47] M. Zeilik, C. Schau, and N. Mattern, Conceptual astronomy. II. Replicating conceptual gains, probing attitude changes across three semesters, *Am. J. Phys.* **67**, 923 (1999).
- [48] M. Zeilik, C. Schau, N. Mattern, S. Hall, K. W. Teague, and W. J. Bisard, Conceptual astronomy: A novel model for teaching post-secondary science courses, *Am. J. Phys.* **65**, 987 (1997).
- [49] G. Simonelli and C. A. Pilachowski, First-year college students' ideas about astronomy: A pilot study, *Astron. Educ. Rev.* **2**, 166 (2003).
- [50] I. Heyer, S. J. Slater, and T. Slater, Establishing the empirical relationship between non-science majoring undergraduate learners' spatial thinking skills and their conceptual astronomy knowledge, *Rev. Latinoam. Educ. Astron.* **45** (2013).
- [51] S. J. Slater, S. P. Schleigh, and D. J. Stork, Analysis of individual Test Of Astronomy Standards (TOAST) item responses, *J. Astron. Earth Sci. Educ.* **2**, 89 (2015).
- [52] S. J. Slater, The development and validation of the Test Of Astronomy STandards (TOAST), *J. Astron. Earth Sci. Educ.* **1**, 1 (2014).
- [53] D. Pundak, I. Liberman, and M. Shacham, From conceptual frameworks to mental models for astronomy: Students' perceptions, *J. Astron. Earth Sci. Educ.* **4**, 109 (2017).
- [54] H. Sugarman, C. Impey, S. Buxner, and J. Antonellis, Astrology beliefs among undergraduate students, *Astron. Educ. Rev.* **10**, 010101 (2011).
- [55] J. Blue, Assessing astronomy students' views about the nature of scientific inquiry, *Phys. Rev. Phys. Educ. Res.* **14**, 010141 (2018).
- [56] J. M. Bailey, Measuring students' understanding in astronomy with research-based assessment tools, in *Astronomy Education: Evidence-Based Instruction for Introductory Courses*, edited by S. Buxner and C. Impey (IOP Publishing, Bristol, United Kingdom, 2020), p. 10.
- [57] J. M. Bailey, Concept inventories for ASTRO 101, *Phys. Teach.* **47**, 439 (2009).
- [58] I. A. Halloun and D. Hestenes, The initial knowledge state of college physics students, *Am. J. Phys.* **53**, 1043 (1985).
- [59] D. Hestenes, M. Wells, and G. Swackhamer, Force concept inventory, *Phys. Teach.* **30**, 141 (1992).
- [60] R. S. Lindell and J. P. Olsen, Developing the lunar phases concept inventory, in *Proceedings of the 2002 Physics Education Research Conference, Boise, Idaho*, edited by S. V. Franklin, K. Cummings, and J. Marx (AAPT PERTG, College Park, MD, 2002).
- [61] J. M. Bailey, Development of a concept inventory to assess students' understanding and reasoning difficulties about the properties and formation of stars, *Astron. Educ. Rev.* **6**, 133 (2007).
- [62] J. M. Bailey, B. Johnson, E. E. Prather, and T. F. Slater, Development and validation of the star properties concept inventory, *Int. J. Sci. Educ.* **34**, 2257 (2012).
- [63] C. S. Wallace and J. M. Bailey, Do concept inventories actually measure anything?, *Astron. Educ. Rev.* **9**, 010116 (2010).
- [64] E. M. Bardar, E. E. Prather, K. Brecher, and T. F. Slater, Development and validation of the light and spectroscopy concept inventory, *Astron. Educ. Rev.* **5**, 103 (2007).
- [65] W. M. Schlingman, E. E. Prather, C. S. Wallace, A. L. Rudolph, and G. Brissenden, A classical test theory analysis of the light and spectroscopy concept inventory national study data set, *Astron. Educ. Rev.* **11**, 010107 (2012).
- [66] C. S. Wallace, T. G. Chambers, and E. E. Prather, Item response theory evaluation of the light and spectroscopy concept inventory national data set, *Phys. Rev. Phys. Educ. Res.* **14**, 010149 (2018).
- [67] K. E. Williamson, S. D. Willoughby, and E. E. Prather, Development of the Newtonian gravity concept inventory, *Astron. Educ. Rev.* **12**, 010107 (2013).
- [68] K. E. K. Nottis, E. N. Ladd, A. Goodman, and P. Udomprasert, Initial development of a concept inventory

- to assess size, scale, and structure in introductory astronomy, *U.S. China Educ. Rev. B* **5**, 689 (2015).
- [69] M. N. Simon, E. E. Prather, S. R. Buxner, and C. D. Impey, The development and validation of the planet formation concept inventory, *Int. J. Sci. Educ.* **41**, 2448 (2019).
- [70] M. C. LoPresto and J. Hubble-Zdanowski, A life in the Universe survey, *Astron. Educ. Rev.* **11**, 010110 (2012).
- [71] C. S. Wallace, E. E. Prather, and D. K. Duncan, A study of general education astronomy students' understandings of cosmology. Part I. Development and validation of four conceptual cosmology surveys, *Astron. Educ. Rev.* **10**, 010106 (2011).
- [72] C. S. Wallace, E. E. Prather, and D. K. Duncan, A study of general education astronomy students' understandings of cosmology. Part II. Evaluating four conceptual cosmology surveys: A classical test theory approach, *Astron. Educ. Rev.* **10**, 010107 (2011).
- [73] C. S. Wallace, E. E. Prather, and D. K. Duncan, A study of general education astronomy students' understandings of cosmology. Part III. Evaluating four conceptual cosmology surveys: An item response theory approach, *Astron. Educ. Rev.* **11**, 010103 (2012).
- [74] H. Bekaert, H. Van Winckel, W. Van Dooren, A. Steegen, and M. De Cock, Design and validation of an instrument to test students' understanding of the apparent motion of the Sun and stars, *Phys. Rev. Phys. Educ. Res.* **16**, 020135 (2020).
- [75] A. Favia, N. F. Comins, G. L. Thorpe, and D. J. Batuski, A direct examination of college student misconceptions in astronomy: A new instrument, *J. Rev. Astron. Educ. Outreach* **1**, A21 (2014).
- [76] A. Favia, N. F. Comins, G. L. Thorpe, and D. J. Batuski, A direct examination of college student misconceptions in astronomy II: Validity of the Astronomy Beliefs Inventory, *J. Rev. Astron. Educ. Outreach* **1**, A3 (2015).
- [77] D. Pundak, Evaluation of conceptual frameworks in astronomy, *Probl. Educ. 21st Century* **69**, 57 (2016).
- [78] M. C. LoPresto, Astronomy diagnostic test results reflect course goals and show room for improvement, *Astron. Educ. Rev.* **5**, 16 (2006).
- [79] E. Brogt, D. Sabers, E. E. Prather, G. L. Deming, B. Hufnagel, and T. F. Slater, Analysis of the astronomy diagnostic test, *Astron. Educ. Rev.* **6**, 25 (2007).
- [80] E. M. Bardar, First results from the light and spectroscopy concept inventory, *Astron. Educ. Rev.* **6**, 75 (2007).
- [81] R. Freed, D. H. McKinnon, M. Fitzgerald, and C. M. Norris, Development and validation of an astronomy self-efficacy instrument for understanding and doing, *Phys. Rev. Phys. Educ. Res.* **18**, 010117 (2022).
- [82] R. Freed, D. H. McKinnon, M. T. Fitzgerald, and S. Salimpour, Confirmatory factor analysis of two self-efficacy scales for astronomy understanding and robotic telescope use, *Phys. Rev. Phys. Educ. Res.* **19**, 020164 (2023).
- [83] M. M. Wooten, A. M. Cool, E. E. Prather, and K. D. Tanner, Comparison of performance on multiple-choice questions and open-ended questions in an introductory astronomy laboratory, *Phys. Rev. ST Phys. Educ. Res.* **10**, 020103 (2014).
- [84] T. H. Robertson and W. H. Finch, Assessment of large general education astronomy classes, *Astron. Educ. Rev.* **4**, 28 (2006).
- [85] Z. Richards and A. M. Kelly, Predicting community college astronomy performance through logistic regression, *Phys. Rev. Phys. Educ. Res.* **19**, 010119 (2023).
- [86] J. M. Bailey, D. Lombardi, J. R. Cordova, and G. M. Sinatra, Meeting students halfway: Increasing self-efficacy and promoting knowledge change in astronomy, *Phys. Rev. Phys. Educ. Res.* **13**, 020140 (2017).
- [87] K. J. Berryhill and T. F. Slater, Opportunity to learn: Investigating possible predictors for pre-course Test Of Astronomy STandards TOAST scores, *J. Astron. Earth Sci. Educ.* **4**, 95 (2017).
- [88] S. R. Buxner, C. D. Impey, J. Romine, and M. Nieberding, Linking introductory astronomy students' basic science knowledge, beliefs, attitudes, sources of information, and information literacy, *Phys. Rev. Phys. Educ. Res.* **14**, 010142 (2018).
- [89] C. S. Wallace, E. E. Prather, and B. M. Mendelsohn, ASTRO 101 students' perceptions of science: Results from the thinking about science survey instrument, *Astron. Educ. Rev.* **12**, 010101 (2013).
- [90] W. W. Cobern, *The Thinking about Science Survey Instrument (TSSI)—SLCSP 151*, Scientific Literacy and Cultural Studies Project (Western Michigan University, Kalamazoo, MI, 2001).
- [91] Z. Bergstrom, P. M. Sadler, and G. Sonnert, Evolution and persistence of students' astronomy career interests: A gender study, *J. Astron. Earth Sci. Educ.* **3**, 77 (2016).
- [92] D. Lombardi, T. F. Shipley, Astronomy Team, Chemistry Team, Engineering Team, Geoscience Team, and Physics Team, The curious construct of active learning, *Psychol. Sci. Publ. Interest* **22**, 8 (2021).
- [93] S. Freeman, S. L. Eddy, M. McDonough, M. K. Smith, N. Okoroafor, H. Jordt, and M. P. Wenderoth, Active learning increases student performance in science, engineering, and mathematics, *Proc. Natl. Acad. Sci. U.S.A.* **111**, 8410 (2014).
- [94] E. E. Prather, A. L. Rudolph, and G. Brissenden, Teaching and learning astronomy in the 21st century, *Phys. Today* **62**, No. 10, 41 (2009).
- [95] C. J. Ballen, C. Wieman, S. Salehi, J. B. Searle, and K. R. Zamudio, Enhancing diversity in undergraduate science: Self-efficacy drives performance gains with active learning, *CBE Life Sci. Educ.* **16**, ar56 (2017).
- [96] A. L. Rudolph, E. E. Prather, G. Brissenden, D. Consiglio, and V. Gonzaga, A national study assessing the teaching and learning of introductory astronomy part II: The connection between student demographics and learning, *Astron. Educ. Rev.* **9**, 010107 (2010).
- [97] E. E. Prather, T. F. Slater, J. P. Adams, J. M. Bailey, L. V. Jones, and J. A. Dostal, Research on a Lecture-Tutorial approach to teaching introductory astronomy for non-science majors, *Astron. Educ. Rev.* **3**, 122 (2004).
- [98] E. E. Prather and C. S. Wallace, Lecture-Tutorials in introductory astronomy, in *Astronomy Education: Evidence-Based Instruction for Introductory Courses*, edited by C. Impey and S. Buxner (IOP Publishing, Bristol, United Kingdom, 2020), p. 3.

- [99] W. R. Alexander, Assessment of teaching approaches in an introductory astronomy college classroom, *Astron. Educ. Rev.* **3**, 178 (2005).
- [100] J. P. Adams, E. E. Prather, and T. F. Slater, *Lecture-Tutorials for Introductory Astronomy*, preliminary ed. (Prentice Hall, Upper Saddle River, NJ, 2002).
- [101] M. C. LoPresto and S. R. Murrell, Using the star properties concept inventory to compare instruction with lecture tutorials to traditional lectures, *Astron. Educ. Rev.* **8**, 010105 (2009).
- [102] J. W. Eckenrode, E. E. Prather, and C. S. Wallace, Correlations between students' written responses to Lecture-Tutorial questions and their understandings of key astrophysics concepts, *J. Coll. Sci. Teach.* **45**, 65 (2016).
- [103] E. E. Prather, A. L. Rudolph, G. Brissenden, and W. M. Schlingman, A national study assessing the teaching and learning of astronomy. Part I. The effect of interactive instruction, *Am. J. Phys.* **77**, 320 (2009).
- [104] M. C. LoPresto, Using visual assessments and tutorials to teach solar system concepts in introductory astronomy, *Astron. Educ. Rev.* **9**, 010101 (2010).
- [105] C. S. Wallace, E. E. Prather, and D. K. Duncan, A study of general education astronomy students' understandings of cosmology. Part V. The effects of a new suite of cosmology lecture-tutorials on students' conceptual knowledge, *Int. J. Sci. Educ.* **34**, 1297 (2012).
- [106] C. S. Wallace, C. Hatcher, T. G. Chambers, S. D. Hornstein, J. Kamenetzky, and E. E. Prather, A new lecture-tutorial for teaching interferometry to ASTRO 101 students, *Phys. Teach.* **59**, 440 (2021).
- [107] C. S. Wallace, E. E. Prather, S. D. Hornstein, J. O. Burns, W. M. Schlingman, and T. G. Chambers, A new Lecture-Tutorial for teaching about molecular excitations and synchrotron radiation, *Phys. Teach.* **54**, 40 (2016).
- [108] M. C. LoPresto and T. F. Slater, A new comparison of active learning strategies to traditional lectures for teaching college astronomy, *J. Astron. Earth Sci. Educ.* **3**, 59 (2016).
- [109] D. P. Maloney, Ranking tasks: A new type of test item, *J. Coll. Sci. Teach.* **16**, 510 (1987), <https://www.jstor.org/stable/42987297>.
- [110] D. W. Hudgins, E. E. Prather, D. J. Grayson, and D. P. Smits, Effectiveness of collaborative ranking tasks on student understanding of key astronomy concepts, *Astron. Educ. Rev.* **5**, 1 (2006).
- [111] P. J. Green, *Peer Instruction for Astronomy*, Prentice Hall Series in Educational Innovation (Prentice Hall, Upper Saddle River, NJ, 2003).
- [112] F. Lyman, The responsive classroom discussion: The inclusion of all students, in *Mainstreaming Digest*, edited by A. S. Anderson (University of Maryland College of Education, College Park, MD, 1981), p. 109.
- [113] A. Crider, Hot seat questioning: A technique to promote and evaluate student dialogue, *Astron. Educ. Rev.* **3**, 137 (2004).
- [114] P. M. Len, Different reward structures to motivate student interaction with electronic response systems in astronomy, *Astron. Educ. Rev.* **5**, 5 (2007).
- [115] M. C. James, F. Barbieri, and P. Garcia, What are they talking about? Lessons learned from a study of peer instruction, *Astron. Educ. Rev.* **7**, 37 (2008).
- [116] E. E. Prather and G. Brissenden, Clickers as data gathering tools and students' attitudes, motivations, and beliefs on their use in this application, *Astron. Educ. Rev.* **8**, 010103 (2009).
- [117] C. H. Crouch and E. Mazur, Peer instruction: Ten years of experience and results, *Am. J. Phys.* **69**, 970 (2001).
- [118] A. L. Rudolph, Do you always need a textbook to teach ASTRO 101?, *Astron. Educ. Rev.* **12**, 010112 (2013).
- [119] Survey of attitudes toward astronomy, field-tested learning assessment guide website: <http://www.flaguide.org/tools/attitude/astpr.php> [accessed June 8, 2004].
- [120] W. J. Straits and R. R. Wilke, Activities-based astronomy: An evaluation of an instructor's first attempt and its impact on student characteristics, *Astron. Educ. Rev.* **2**, 46 (2003).
- [121] G. L. Deming, S. T. Miller, and J. D. Trasco, *Cooperative Learning Activities in Introductory Astronomy for Non-Science Majors* (University of Maryland, College Park, MD, 1997).
- [122] M. Zeilik and V. J. Morris, The impact of cooperative quizzes in a large introductory astronomy course for non-science majors, *Astron. Educ. Rev.* **3**, 51 (2004).
- [123] S. T. Miller and C. R. James, What is the impact of collaborative exams on learning and attitudes in introductory astronomy classes?, *J. Astron. Earth Sci. Educ.* **6**, 17 (2019), <https://clutejournals.com/index.php/JAESE/article/view/10289>.
- [124] S. T. Miller and C. R. James, The effect of animations within PowerPoint presentations on learning introductory astronomy, *Astron. Educ. Rev.* **10**, 010202 (2011).
- [125] D. K. Duncan, A. R. Hoekstra, and B. R. Wilcox, Digital devices, distraction, and student performance: Does in-class cell phone use reduce learning?, *Astron. Educ. Rev.* **11**, 010108 (2012).
- [126] K. C. Yu, K. Sahami, V. Sahami, and L. C. Sessions, Using a digital planetarium for teaching seasons to undergraduates, *J. Astron. Earth Sci. Educ.* **2**, 33 (2015).
- [127] K. C. Yu, K. Sahami, G. Denn, V. Sahami, and L. C. Sessions, Immersive planetarium visualizations for teaching solar system Moon concepts to undergraduates, *J. Astron. Earth Sci. Educ.* **3**, 93 (2016).
- [128] K. C. Yu, K. Sahami, and J. Dove, Learning about the scale of the solar system using digital planetarium visualizations, *Am. J. Phys.* **85**, 550 (2017).
- [129] N. Fajriani and M. Masturi, Utilization of sky map application in astronomy learning of celestial coordinates to improve students' understanding of concepts and digital literacy for physics education students, *Berkala Ilm. Pendidik. Fis.* **11**, 80 (2023).
- [130] D. K. Duncan and L. Arthurs, Improving student attitudes about learning science and student scientific reasoning skills, *Astron. Educ. Rev.* **11**, 010102 (2012).
- [131] H. M. L. G. Flohic, Comparing the impact of an astronomy course and a science and society seminar on undergraduate students' attitudes toward science, *J. Astron. Earth Sci. Educ.* **4**, 75 (2017).

- [132] A. Elby, Helping physics students learn how to learn, *Am. J. Phys.* **69**, S54 (2001).
- [133] C. A. Garland and D. L. Ratay, Using literacy techniques to teach astronomy to non-science majors, *Astron. Educ. Rev.* **6**, 59 (2007).
- [134] E. M. Bardar and K. Brecher, Project LITE educational materials and their effectiveness as measured by the light and spectroscopy concept inventory, *Astron. Educ. Rev.* **6**, 85 (2008).
- [135] A. O. Meyer, M. J. Mon, and S. T. Hibbard, The lunar phases project: A mental model-based observational project for undergraduate nonscience majors, *Astron. Educ. Rev.* **10**, 010203 (2011).
- [136] M. L. Allen and D. Kelly-Riley, Promoting undergraduate critical thinking in ASTRO 101 lab exercises, *Astron. Educ. Rev.* **4**, 10 (2005).
- [137] I. C. Jacobi, H. J. Newberg, D. Broder, R. A. Finn, A. J. Milano, L. A. Newberg, A. T. Weatherwax, and D. C. B. Whittet, Effect of night laboratories on learning objectives for a nonmajor astronomy class, *Astron. Educ. Rev.* **7**, 66 (2008).
- [138] K. J. Sibbersen, Impact of collaborative groups versus individuals in undergraduate inquiry-based astronomy laboratory learning exercises, *J. Astron. Earth Sci. Educ.* **1**, 41 (2014).
- [139] R. S. Schwartz, N. G. Lederman, and J. S. Lederman, An instrument to assess views of scientific inquiry: The VOSI questionnaire, in *Proceedings of the 2008 Annual Meeting of the National Association for Research in Science Teaching* (NARST, Baltimore, MD, 2008).
- [140] X. C. Cid and R. E. Lopez, The impact of stereo display on student understanding of phases of the Moon, *Astron. Educ. Rev.* **9**, 010105 (2010).
- [141] N. Ladd, K. E. K. Nottis, and P. Udomprasert, A hybrid hands-on and computer simulation laboratory activity for the teaching of astronomical parallax, *J. Astron. Earth Sci. Educ.* **6**, 31 (2019).
- [142] T. F. Slater and L. V. Jones, Assessment of an internet-delivered interactive approach to introductory astronomy for non-science majors, *Astron. Educ. Rev.* **3**, 17 (2004).
- [143] S. T. Miller and S. L. Redman, Enhancing student performance in an online introductory astronomy course with video demonstrations, *Astron. Educ. Rev.* **9**, 010114 (2010).
- [144] S. T. Miller and S. L. Redman, Improving instructor presence in an online introductory astronomy course through video demonstrations, *Astron. Educ. Rev.* **9**, 010115 (2010).
- [145] S. T. Miller, Increasing student participation in online group discussions via Facebook, *Astron. Educ. Rev.* **12**, 010103 (2013).
- [146] D. F. Barringer, J. D. Plummer, J. Kregenow, and C. Palma, Gamified approach to teaching introductory astronomy online, *Phys. Rev. Phys. Educ. Res.* **14**, 010140 (2018).
- [147] L. B. Horodyskyj, C. Mead, Z. Belinson, S. Buxner, S. Semken, and A. D. Anbar, Habitable worlds: Delivering on the promises of online education, *Astrobiology* **18**, 86 (2018).
- [148] J. J. Trout, T. Jacobsen, G. Buondonno, and C. Weber, “Astrophysics: History and theory”, an undergraduate, interdisciplinary, asynchronous astrophysics course taught in the physics program, *Phys. Educ.* **57**, 045015 (2022).
- [149] A. Fraknoi, D. Morrison, and S. C. Wolff, *Astronomy* (OpenStax, Houston, TX, 2022).
- [150] American Association for the Advancement of Science, *Science for all Americans* (Oxford University Press, New York, 1990).
- [151] National Research Council, *National Science Education Standards* (National Academy Press, Washington, DC, 1996).
- [152] H. B. Hewitt, M. N. Simon, C. Mead, S. Grayson, G. L. Beall, R. T. Zelle, K. Tock, and K. A. Pearson, Development and assessment of a course-based undergraduate research experience for online astronomy majors, *Phys. Rev. Phys. Educ. Res.* **19**, 020156 (2023).
- [153] M. M. Wooten, K. Coble, A. W. Puckett, and T. Rector, Investigating introductory astronomy students’ perceived impacts from participation in course-based undergraduate research experiences, *Phys. Rev. Phys. Educ. Res.* **14**, 010151 (2018).
- [154] C. H. S. Lacy, Observational research for all students, *Astron. Educ. Rev.* **2**, 129 (2003).
- [155] A. Fraknoi, Insights from a survey of astronomy instructors in community and other teaching-oriented colleges in the United States, *Astron. Educ. Rev.* **3**, 7 (2004).
- [156] N. Lippert and B. Partridge, To hear ourselves as others hear us, *Astron. Educ. Rev.* **3**, 31 (2004).
- [157] E. Brogt, Instruction as a scientific experiment: A professional development case study of a professor changing the introductory astronomy course for non-science majors, *Astron. Educ. Rev.* **6**, 20 (2007).
- [158] J. M. Bailey and K. Nagamine, Experiencing conceptual change about teaching: A case study from astronomy, *Am. J. Phys.* **80**, 542 (2012).
- [159] J. A. Dole and G. M. Sinatra, Reconceptualizing change in the cognitive construction of knowledge, *Educ. Psychol.* **33**, 109 (1998).
- [160] D. J. Everding and J. M. Keller, Survey of the academic use of planetariums for undergraduate education, *Phys. Rev. Phys. Educ. Res.* **16**, 020128 (2020).
- [161] D. J. Everding and J. M. Keller, Study of faculty instructors in undergraduate classroom and planetarium learning environments, *J. Astron. Earth Sci. Educ.* **8**, 39 (2021).
- [162] J. D. Plummer and K. J. Small, Informal science educators’ pedagogical choices and goals for learners: The case of planetarium professionals, *Astron. Educ. Rev.* **12**, 010105 (2013).
- [163] K. J. Small and J. D. Plummer, Survey of the goals and beliefs of planetarium professionals regarding program design, *Astron. Educ. Rev.* **9**, 010112 (2010).
- [164] National Research Council, *How People Learn: Brain, Mind, Experience, and School: Expanded Edition* (The National Academies Press, Washington, DC, 2000).
- [165] National Research Council, *How Students Learn: History, Mathematics, and Science in the Classroom* (The National Academies Press, Washington, DC, 2005).

- [166] S. K. Schultz and T. F. Slater, Who are the planetarians? A demographic survey of planetarium-based astronomy educators, *J. Astron. Earth Sci. Educ.* **7**, 25 (2020).
- [167] S. K. Schultz and T. F. Slater, Use of formative assessment-based active learning by astronomy educators teaching in live planetarium learning environments, *J. Astron. Earth Sci. Educ.* **8**, 27 (2021).
- [168] M. L. Berendsen, Conceptual astronomy knowledge among amateur astronomers, *Astron. Educ. Rev.* **4**, 1 (2005).
- [169] V. Yocco, E. C. Jones, and M. Storksdieck, Factors contributing to amateur astronomers' involvement in education and public outreach, *Astron. Educ. Rev.* **11**, 010109 (2012).
- [170] A. Price, P. L. Gay, T. Searle, and G. Brissenden, A history and informal assessment of the slacker astronomy podcast, *Astron. Educ. Rev.* **5**, 53 (2006).
- [171] M. J. Raddick, G. Bracey, P. L. Gay, C. J. Lintott, C. Cardamone, P. Murray, K. Schawinski, A. S. Szalay, and J. Vandenberg, Galaxy Zoo: Motivations of citizen scientists, *Astron. Educ. Rev.* **12**, 010106 (2013).
- [172] E. E. Prather, S. Cormier, C. S. Wallace, C. Lintott, M. J. Raddick, and A. Smith, Measuring the conceptual understandings of citizen scientists participating in Zooniverse projects: A first approach, *Astron. Educ. Rev.* **12**, 010109 (2013).
- [173] M. J. Raddick, E. E. Prather, and C. S. Wallace, Galaxy Zoo: Science content knowledge of citizen scientists, *Publ. Understanding Sci.* **28**, 636 (2019).
- [174] C. Chen, G. Sonnert, P. M. Sadler, D. Sassselov, and C. Fredericks, The impact of student misconceptions on student persistence in a MOOC, *J. Res. Sci. Teach.* **57**, 879 (2020).
- [175] C. Impey and M. Formanek, MOOCS and 100 days of COVID: Enrollment surges in massive open online astronomy classes during the coronavirus pandemic, *Soc. Sci. Hum. Open* **4**, 100177 (2021).
- [176] C. Impey, M. Formanek, S. Buxner, and M. Wenger, Science knowledge and attitudes of lifelong learners in an astronomy massive open online course, *Int. J. Sci. Educ.* **11**, 110 (2021).
- [177] S. Salimpour, M. T. Fitzgerald, and U. Eriksson, A descriptive overview of English-language publications in the field of astronomy education research, 1898 to 2022, *Astron. Educ. J.* **4**, 140aer (2024).
- [178] T. Campbell, C. Schwarz, and M. Windschitl, What we call misconceptions may be necessary stepping-stones toward making sense of the world, *Sci. Teach.* **83**, 69 (2016).
- [179] D. Lombardi and J. M. Bailey, Science teaching and learning, in *Handbook of Educational Psychology*, edited by P. Schutz and K. R. Muis (Routledge, New York, NY, 2024), p. 531, [10.4324/9780429433726-27](https://doi.org/10.4324/9780429433726-27).
- [180] J. Henrich, S. J. Heine, and A. Norenzayan, The weirdest people in the world?, *Behav. Brain Sci.* **33**, 61 (2010).
- [181] E. L. Gomez and M. T. Fitzgerald, Robotic telescopes in education, *Astron. Rev.* **13**, 28 (2017).
- [182] D. E. Reichart, Robotic telescope labs for survey-level undergraduates, *Phys. Teach.* **59**, 728 (2021).
- [183] R. Freed, D. H. McKinnon, S. Salimpour, M. Fitzgerald, D. Reichart, and C. Norris, Self-efficacy changes and gender effects on self-efficacy in a large-scale robotic telescope focused curriculum, *Phys. Rev. Phys. Educ. Res.* **20**, 010137 (2024).
- [184] J. P. Byrnes, *Cognitive Development for Academic Achievement: Building Skills and Motivation* (The Guilford Press, New York, NY, 2021).
- [185] R. S. French, Everyone's Universe: Teaching astronomy in community colleges, in *Astronomy Education: Evidence-Based Instruction for Introductory Courses*, edited by S. Buxner and C. Impey (IOP Publishing, Bristol, United Kingdom, 2020), p. 11.