IAU ASTRONOMY EDUCATION AND DEVELOPMENT
Division C - C1 Commission

AEJ

ASTRONOMY EDUCATION JOURNAL

Volume 1
Issue 1
2021

ISSN: 2004-2981

Published in Sweden • Urban Eriksson • Paulo S. Bretones

Astronomy Education Journal, 2021, Vol. 01, No. 1

doi: 10.32374/AEJ.2021.1.1.001

ISSN: 2004-2981 · Published in Sweden

EDITORIAL

Welcome to AEJ

Urban Eriksson[1,*] and Paulo S. Bretones[2]

[1]Lund University, Sweden ; [2]Universidade Federal de São Carlos, Brazil

[*]urban.eriksson@fysik.lu.se

1 Introduction

Dear Colleagues,

We are pleased to launch for the general public, in particular those with interests in research and teaching of Astronomy, the first issue of the Astronomy Education Journal (AEJ).

The AEJ was announced with the first call for papers in 2nd December 2020. Since then, there were 35 submitted papers, 9 papers were rejected, and 10 papers are yet in the review process.

In this issue we publish seven articles in two sections:

- Astronomy Education Research
- Astronomy Education and Practice

2 Astronomy Education Research

- Assessing preservice elementary teachers' enactment of science practices using children's astronomy storybooks, by *Julia D. Plummer, Kyungjin Cho, Christopher Palma, Daniel F. Barringer, Timothy Gleason and Katie Nolan.*
- Interactive cosmology visualization using the Hubble Ultra Deep Field data in the classroom, by *Liam J. Nolan, Mira R. Mechtley, Rogier A.Windhorst, Karen Knierman, Teresa A. Ashcraft, Seth H. Cohen, Scott Tompkins and Lisa M.Will*

3 Astronomy Education and Practice

- A historical method approach to teaching Kepler's 2nd law, by *Wladimir Lyra.*
- An observational project for a large class – determination of the duration of the sidereal day, by *William Tobin.*
- Timing pulsars: An exercise in statistical analysis and the scientific process, by *John R. Walkup, Joseph White and Roger Key.*
- Resource guides for astronomy educators and their students, by *Andrew Fraknoi.*
- Should we teach general relativity in high school? Why and how?, by *Jorge. E. Horvath and Pedro. H. R. S. Moraes*

4 Background

At the inaugural IAU Astronomy Education Conference (AstroEdu) in 2019, we announced the intention to launch a new Astronomy Education Journal (AEJ). Today, we are happy to inform you that the journal is publishing its first issue. This online journal aims to be a key global publication platform for both researchers and practitioners, in the field of Astronomy Education, Research, and Methods.

AEJ aims to meet the needs of the astronomy education community by providing a location for all manner of practical, newsworthy and scholarly publications involving developments in the field. In a sense, the journal tries to capture the original spirit whilst taking on board the important lessons from the, now out-of-print, Astronomy Education Review.

By focusing on building community collaboration, disseminating important news and opinions, while also maintaining a section on more formal, technical, Astronomy Education Research (AER). This research section intends to compliment the current scholarly discipline-based work undertaken by, for example, Latin-American Journal of Astronomy Education (RELEA), the Journal of Astronomy & Earth Sciences Education (JAESE) and, recently, the acceptance of AER articles into Physical Review Physics Education Research (PRPER).

Inspired by our sibling, IAU Commission C2: Communicating Astronomy with the Public journal, the CAP journal, we will accept various types of articles. AEJ will draw on journals such as the CAP Journal, Nature, and Science, to incorporate both peer-reviewed and non-peer reviewed articles. There will be a peer-reviewed section of research articles that will be incorporated into AEJ's scholarly indices. These research articles will be formally peer-reviewed as traditional scientific journal manuscripts and, as such, need to be of a sufficient scholarly standard as recommended by, for example, Scopus. In addition, there is also scope for published invited reviews written by specialists of the area of AER. There will also be a less formal, non-peer-reviewed, but edited and curated section that contains other relevant material, such as, news, announcements, interviews, opinions, resources, correspondences, best-practices, classroom and astronomical activities, to help circulate information among the community.

More information about the journal and instructions for authors can be found at: `https://www.astroedjournal.org`. We welcome everyone to submit manuscripts to AEJ.

You can use the text above to advertise this release of AEJ in your countries or international platforms to reach the interested readers and authors.

5 Acknowledgements

We are grateful to the managing editor Dr. Saeed Salimpour for his work towards the publication of this issue, the editorial board, authors, referees, and all those who, directly or indirectly, assists us in the continuity of this initiative and, in particular, in the preparation of this edition.

Astronomy Education Journal, 2021, Vol. 01, No. 1

doi: 10.32374/AEJ.2021.1.1.002

RESEARCH ARTICLE

Assessing preservice elementary teachers' enactment of science practices using children's astronomy storybooks

Julia D. Plummer[1,*], Kyungjin Cho[2], Christopher Palma[1], Daniel F. Barringer[3], Timothy Gleason[1] and Katie Nolan[1]

[1]The Pennsylvania State University, USA ; [2]Duksung Women's University, South Korea ; [3]Texas State University, USA

[*]jdp17@psu.edu

Abstract

Purposefully-designed, science content courses have the potential to help prepare future elementary teachers by helping to develop their understanding of the practices of science. We extend research on how science content courses can prepare future elementary teachers by investigating how students' experiences in such a course contributes to their ability to enact coherent use of science practices within a science investigation. We investigated U.S. college students' enactment of coherent science inquiry investigations after completing an inquiry-based astronomy course designed for preservice elementary teachers. We assessed preservice teachers' (N=63) enactment of the coherence between question, data gathering, and evidence-based explanations using a novel format: student-generated astronomy-based children's storybooks. Most students (59%) wrote storybooks featuring coherent investigations; in other words, their stories featured characters who linked an investigation question to data collection and to an evidence-based explanation. Over the three years of data collection, the percentage of preservice teachers who wrote coherent investigations in their final storybooks increased from 35% to 71% suggesting that additional scaffolding provided by the faculty in years 2 and 3 helped students understand these practices. Our findings suggest that purposefully-designed, science content courses can help preservice teachers learn about coherent science inquiry in astronomy. We also suggest that projects tied to the students' own future careers, such as creating children's science storybooks, can be used as assessment tools by faculty to assess preservice teachers' development of science practices.

Keywords: Science practices; Preservice teachers; Storybooks

1 Introduction

The *Next Generation Science Standards* (NGSS Lead States, 2013) have raised the bar for what is expected of preservice elementary teachers in the U.S. as they begin their careers teaching science to young children. The current reform-based approach to teaching science is likely to challenge new elementary teachers in the complex ways it represents science practices (McNeill et al., 2017); their own K-12 experiences, as well as experiences in college science classrooms, frequently present science in a more fragmented, fact-oriented view, rather than one that embraces questions, evidence, and inquiry (National Research

Council, 2012; Roth and Garnier, 2007). Yet, university-level science courses designed to support this reform-based perspective can make a difference in preparing new elementary teachers for future science teaching (Avraamidou and Zembal-Saul, 2010; Haefner and Zembal-Saul, 2004). This study investigated one such university science course for preservice elementary teachers and how a novel assessment opportunity—writing a children's storybook—might provide insights into these future teachers' understanding of science practices.

2 Conceptual Framework: Coherent Science Inquiry Investigations

Our research highlights the importance of preservice teachers' developing an understanding of science practices. Science practices describe behaviours that scientists engage in as they hypothesize and investigate about the natural world. The science practices framework we draw on is based on the *coherent science inquiry investigation* (CSII)(Plummer and Tanis Ozcelik, 2015), which focuses on the coherence between scientific questions, data gathering, and evidence-based explanations during investigations. Plummer and Tanis Ozcelik Plummer and Tanis Ozcelik adapted the CSII framework from Roth and colleague's *coherent science content storyline* (CSCS) (Roth and Garnier, 2007; Roth et al., 2011). Roth and colleagues recommend that teachers develop lessons that follow CSCS, which involves identifying one main learning goal, communicating that goal to students, selecting activities and representations that reflect that goal, and carefully sequencing activities in ways that build toward that goal. Building on the ways Roth et al. used coherence — an inter-connection between goals and activities to form a whole — the CSII organizes students' consistent use of science practices around the investigation's main goal as the unit of analysis (Plummer and Tanis Ozcelik, 2015). In a CSII experience, students' investigation focuses on making sense of a single phenomenon or related set of phenomena as they construct explanations based on evidence in response to a question or problem about the phenomenon; they may also use other science practices, such as modelling or argumentation, during the investigation in ways that support the development of their explanation. Educators design CSIIs by purposefully choosing and sequencing activities in ways that help students attempt to answer the question by making sense of data to form an evidence-based explanation (Plummer and Tanis Ozcelik, 2015). Focusing on coherence between practices towards developing an evidence-based explanation may allow students to develop deeper understandings of the phenomenon and the use of science practices.

An important component of U.S. preservice teachers' professional development is learning to understand and use the NGSS. The NGSS provides science content standards for the physical sciences, life sciences, earth and space sciences, and engineering and technology; standards describing cross-cutting concepts (core ideas that are relevant within and across the different disciplinary areas); and standards for science practices. These standards were developed to improve science education for students in grades K-12 (ages 5-18) in the U.S. by guiding curriculum and assessment development. Many U.S. preservice teachers will be expected to address these standards in their classrooms in the future. Therefore, we drew on recent frameworks for science practices providing guidance and explanation for the Next Generation Science Standards in the U.S. (National Research Council, 2012; McNeill et al., 2017) and providing specific guidance for teachers working with elementary students in science (McNeill, 2011; Zembal-Saul et al., 2013) to provide additional details to our initial CSII framework in order to clarify our definition of specific science practices used in our study.

Scientists observe and investigate the world towards achieving two goals: "(1) to systematically describe the world and (2) to describe and test theories and explanations of how the world words" (National Research Council, 2012, p. 59). Across all investigations with our students, we emphasized the importance of constructing evidence-based explanations, which "focus on a specific question about a phenomenon and construct a how or why account for that phenomenon" (McNeill et al., 2017, p. 207). As our course was designed primarily for preservice elementary teachers, we began with helping them learn to make careful observations that can lead to identifying patterns that need to be explained or further questions to be explored (Zembal-Saul et al., 2013). Thus, students were encouraged to engage in investigations that generated data both relevant to their question and sufficient to support the claims being made about the phenomenon (NGSS Lead States, 2013; Windschitl, 2017). The how or why account (i.e., scientific reasoning) in the evidence-based explanations draws on a science model or science principle to construct a causal account. We engaged our students in further investigation to test explanatory models, as this can help deepen their understanding of the theories that explain their observations and extend their understanding of how scientists investigate the world.

3 Supporting Preservice Teachers' Understanding and Enactment of Science Practices

To support elementary students' conceptual understanding of natural phenomena, elementary teachers need to be prepared to support students in constructing evidence-based explanations during instruction (McNeill, 2011; Zangori and Forbes, 2013). Studies in teacher education indicate that specially designed content courses for preservice elementary teachers may help them develop this understanding of science practices. These courses are designed to prepare teachers for teaching with science practices by engaging them in content investigations that bring particular science practices to the forefront of the discussion (Haefner and Zembal-Saul, 2004). In other words, by considering a sociocultural perspective on learning (Vygotsky, 1986) and a model of learning science practices that is developed through extended teacher participation in learning moments by using those practices (Ford, 2008), preservice teachers are hypothesized to develop a more integrated knowledge of science and science practices, for the purpose of their future teaching, through their own personal actions of engaging in scientific inquiry with their peers.

Studies with preservice and new elementary teachers provide some insight into how a purposefully designed course for preservice teachers might support their development of teaching practices. Haefner and Zembal-Saul (2004) investigated an innovative science content course designed to engage preservice elementary teachers in science inquiry. Through participation in the course, the preservice teachers' views of science shifted towards one that emphasized scientific process over product as they developed an increased understanding of the experimental aspects of science. Further, after participation in the course the participants "became more accepting of approaches to teaching science that encourage children to investigate phenomena about which they have questions" (p. 1670). Avraamidou and Zembal-Saul (2010) investigated two first year teachers' use of science practices with students in their classrooms. As undergraduates, one of the teachers had taken three courses specifically designed to support preservice teachers' understanding of the practices of science; the other had only taken traditional science courses. The teacher who had taken the specially designed science course engaged her students in the language of constructing and communicating claims through inquiry-based investigations while the other new teacher used limited scientific discourse in her classroom. Drawing on evidence from interviews that highlight coherence between the teachers' knowledge of science teaching practices and beliefs about what experiences shaped their teaching, Avraamidou and Zembal-Saul point to the important role played by the purposefully designed content courses in shaping the first teacher's practice. Zangori and Forbes (2013) also emphasized the importance of attending to preservice teachers' own understanding and enactment of science practices as this will shape how they engage students in the future. Using multiple case study design, they found that elementary preservice teachers who had difficulty conceptualizing their own ideas about constructing evidence-based explanations struggled to support their students in this practice.

Research on college-level astronomy courses provides additional insight into how a science content course designed around science practices might help develop preservice elementary teachers' understanding of science practices. Plummer and Tanis Ozcelik (2015) studied preservice teachers who engaged in an extended astronomy investigation that scaffolded students' participation in collaborative sense-making around modelling and evidence-based explanations, as part of their science methods course. Preservice teachers who developed CSIIs in their lesson plans included more complex sense-making practices (i.e., evidence-based explanations and generating representations) than preservice teachers who did not develop for CSIIs in their lessons. Plummer and Tanis Ozcelik also found that preservice teachers with higher astronomy content scores were more likely to develop coherent science investigations in the lesson plans they wrote than students with lower scores. Slater, Slater, and Shaner (2008) investigated a course for preservice elementary teachers where students learned astronomy through *backwards-faded scaffolding* investigations: removing supports to shift towards open-ended inquiry over the course of the semester. After taking the course, students showed significant improvement in their understanding of both science practices and the astronomy content. Further research has also found that non-science majors improve their understanding of inquiry after taking an introductory astronomy course using backwards-faded scaffolding investigations (Lyons, 2011; Sibbernsen, 2014). However, another study using the same curriculum materials and the same research instrument in an introductory astronomy course found no improvement (Stewart, 2013). Stewart's study suggests that the nature of the guided experience—how the instructor engages the students in discourse around the investigations—is critical to shaping learners' knowledge of when and how to use science practices.

Taken together, these studies suggest university science courses may play an important role in preparing preservice teachers to engage in science discourse and enact science practices, in preparation for their future teaching (Haefner and Zembal-Saul, 2004; Lyons, 2011; Plummer and Tanis Ozcelik, 2015; Sibbernsen, 2014; Slater et al., 2008). College science courses that emphasize engagement in the practices of science and that scaffold preservice teachers' experiences over time have the opportunity to develop preservice teachers' understanding of science practices, which in turn may shape how they are prepared to cultivate their future elementary students' understanding of science phenomena (Avraamidou and Zembal-Saul, 2010; Zangori and Forbes, 2013).

4 Science Curriculum and Children's Science Storybooks

Our study considered how we can further preservice teachers' education by facilitating their engagement with curriculum materials as we investigated their development of storybooks as a form of curriculum design. Curriculum materials serve as tools to mediate a teacher's planning and enactment in the classroom (Brown, 2009; Forbes, 2011). Teachers need curriculum materials to be accurate with coherent contents, have clear purpose for learning, and provide multiple opportunities for students to represent their ideas (Davis and Krajcik, 2005). However, while elementary teachers often use curriculum materials for their science instruction, they rarely have an opportunity to write their own curriculum materials (Forbes, 2011).

Children's science storybooks are an important classroom resource and could be useful tools to support young learners in engaging with science practices (Plummer and Cho, 2020; Pringle and Lamme, 2005). Murmann and Avraamadou (2014) investigated the use of stories as learning tools for elementary students during an inquiry-based investigation. They found that while the stories have significant potential, their use as a tool for science inquiry is dependent on the teacher's understanding and beliefs about science teaching and learning. This suggested to us that helping our preservice teachers write storybooks that include CSIIs could be an important step towards their own development as teachers who plan to use stories as tools for inquiry. We asked the preservice teachers in our course to write a children's storybook within the domain of astronomy—the focus of our course—that included a character or characters participating in an investigation that concludes with constructing an evidence-based explanation. This type of children's storybook is similar to a *dual-purpose storybook* which not only includes an entertaining, character-driven narrative but also conveys factual science content (Donovan and Smolkin, 2002). While typical dual-purpose children's storybooks present factual information through insets or diagrams, we encouraged students to integrate science content and practices directly through their storybook's narrative. Similar dual-purpose storybooks have been used to support preschool-age children in constructing evidence-based explanations by communicating questions or problems that connect to investigations of scientific phenomena (Plummer and Cho, 2020). Thus, engaging preservice teachers with this storybook format could provide them with a potentially useful pedagogical tool in their future teaching.

The storybooks also served as a novel method to assess our preservice teachers' enactment of science as a process of constructing explanations from evidence through coherent science inquiry investigations. Students needed to take what they learned in our course about astronomy and coherent science inquiry investigations and apply that to writing their storybook. Our study was guided by the following research question: *In what ways do dual-purpose storybooks, written by preservice elementary teachers at the end of an astronomy content course, demonstrate an ability to enact coherent investigations that lead to evidence-based explanations in astronomy?*

5 Methods

5.1 Context of the course

The research subjects were drawn from three separate offerings of an astronomy content course designed specifically for pre-service elementary teachers at a large research university in North-eastern U.S. The course was co-taught by two authors of the present study - faculty in an education department and an astronomy department. The course is organized using a coherent science content storyline format (Roth and Garnier, 2007; Roth et al., 2011) by sequencing investigations towards one main learning goal: gathering evidence in support of Solar System formation as the underlying causal model explaining patterns observed among Solar System objects. We communicated this to the students and carefully selected investigations that built students' understanding towards the larger goal. The course met twice a week for 75 minutes each session; nearly all of the time was dedicated to work on the investigations with very little time spent that would be considered lecture-based. In-class investigations, science notebook assignments, and science report assignments emphasized the coherence of a science inquiry investigation and allowed students to practice writing about connections between a scientific question, data collection methods, and constructing evidence-based explanations. These experiences culminated with students applying what they learned about coherent inquiry investigations in astronomy towards the final assignment, writing a children's science storybooks.

The students worked in small groups to complete a series of seven investigations, called "storyline activities", that address smaller pieces of the storyline (see Table 1). Storyline activities were either guided inquiry (question, background, and procedures provided) or open inquiry (question and background provided) (Buck et al., 2008). In each in-class storyline activity, the instructors provided investigation questions to guide students' engagement in the process of planning their investigation, gathering data, and co-constructing an evidence-based explanation (Zembal-Saul et al., 2013). Students worked collaboratively to investigate astronomy phenomena by collecting data, generating claims, and developing reasoning. Students then participated in whole-class discussions which functioned as a way for small groups to share ideas, critique other groups, and receive feedback from professors and peers on their use of science practices. Students kept daily records throughout the investigations using electronic science notebooks with the web-based software *Evernote*.

Students wrote reports based on three storyline activities (the phases of the Moon, how and why planets orbit the Sun, and the properties of the planets). In the reports, the students summarized their investigation methods, discussed their findings, and included their evidence-based explanations in response to the investigation question provided in class. Though reports were individually written, the students developed their ideas and arguments through collaboration with their small group and whole-class discussions. Professors provided detailed feedback on each student's report to help them improve their understanding of how to communicate the results of a coherent science inquiry investigation.

The storybook assignment was introduced about 8 weeks into the semester. Requirements for the storybooks, as described in the Storybook Assignment Sheet, are included in Table 2. First, the students turned in a short outline that included an investigation question, a description of how the investigation will be carried out in story format, and a description of the final explanation. Each student received written feedback from one of the instructors on their outlines. Near the end of the semester, students were asked to bring a nearly completed version of their books to class for a guided peer review. Using the final grading rubric as a guide, pairs of students provided each other with feedback about the scientific accuracy of the story, the coherence of the investigation, and how their use of images communicated important information about their storybook investigation. Students were able to ask the instructors for additional assistance with the assignment. Final storybooks were submitted in an electronic format to the instructors at the end of the semester. See Figure 1 for example pages from three student storybooks.

5.2 Participants

A total of 67 students were enrolled in the course across the three years of data collection (2015-2017). Almost all students were in their first or second year of college, and had indicated their intent to pursue an elementary education major leading to certification to teach pre-Kindergarten to 4th grade (age 3-11 years). We also collected data from one student pursuing a grade 4-8 certification, one student pursuing secondary education world language certification, one student nearly done with completing her pre-Kindergarten to 4th grade degree, and

Table 1. Storyline activities building towards the formation model of the Solar System

	Storyline activity topic	**Investigation questions**	**Approx. Length**
1	Naked-eye astronomy	How and why does the appearance of the Moon change over time? How does light pollution affect our observations of the stars we observe in the night sky? How does the Sun appear to move and why?	3 weeks
2	Finding planets	Where do we find planets in the sky and how can we use this to predict future planet observations?	2 weeks
3	Investigating orbits	What factors are needed to produce a stable orbit similar to the orbits of planets in our Solar System?	2 weeks
4	How astronomers collect data	What methods can astronomers use to investigate the objects in the Solar System as well as distant stars and why?	2 weeks
5	Craters and the history of the solar system	Do all planet and moon surfaces in the Solar System show the same effects of crater-making impacts from the Late Heavy Bombardment period and why?	2 weeks
6	Solar system properties	How can the planets be grouped according to their properties?	2.5 weeks
7	Formation of the solar system	How does the model of the Solar System's formation explain patterns we observe in the current Solar System?	1.5 weeks

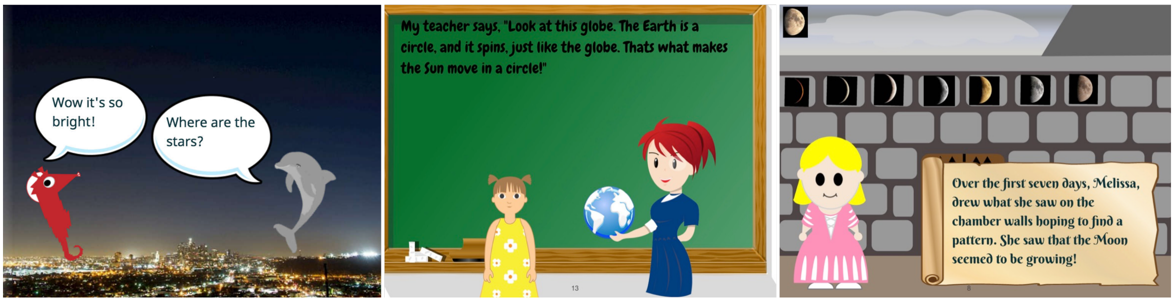

Figure 1. Example storybook pages from "Dolphie and the Stars" by Diane, "Sola and the Sun" by Bethany, and "The Moon Girl" by Ellen.

Table 2. Requirements for Children's Storybook as given in the Storybook Assignment Sheet

Book must contain a creative element that distinguishes this as a children's book as opposed to dry presentation of science fact.
Must include science content but may include a fictional story element (e.g., such as including characters in fictional situations).
Book must be centred around a *Solar System astronomy investigation* but does not have to be one of the specific investigations we did in this class. • Include an investigation question or questions. • Use the story to either encourage the reader to carry out an investigation or show characters in the story carrying out an investigation that answers the investigation question(s). • Demonstrate through your storybook that you are understanding how evidence is used to answer scientific questions. • Demonstrate appropriate methods of collecting data for the astronomy investigation in your storybook. • Integrate into the story the evidence-based explanation that answers the investigation question(s) posed. • Communicate how the reasoning would be constructed for that evidence-based explanation, which may include discussing an appropriate scientific model.
Presentation of the book should reflect elements of a children's book, including the use of photos and/or images that support the narrative, as opposed to an essay or report. • Must include a list of references to ALL images used at the end of the book. • If you want to share this book more broadly in the future, we recommend using public domain images, where possible. Public domain images include those from NASA and Wikimedia Commons.
Book must demonstrate accurate understandings of the science content.
Include a glossary of astronomy terms used, at the end of the book.

one engineering student who did not intend to pursue a teaching certification. Most of the students in the study were female (92%). At the beginning of each semester, students were asked to indicate their willingness to provide their assignments (including storybooks) for this study. All students provided this level of consent for the project. All student names in this manuscript are pseudonyms.

5.3 Data collection

We analysed the final electronic storybooks (N=63) submitted at the end of the semester (some sample attrition occurred due to digital file loss and one student who did not complete the assignment). The number of storybooks gathered for analysis per year was: n=17 in Year 1, n=22 in Year 2, and n=24 in Year 3. The average length of a storybook was 22 pages (SD=8.7) (no length requirement was given in the assignment).

5.4 Analysis

We developed a coding scheme by defining categories based on key elements of the CSIIs represented in the storybooks: use of *investigation questions, planning and carrying out investigations, and constructing explanations* (Table 3). Initially, codes were defined using literature describing these science practices (*e.g.* McNeill et al. 2017; National Research Council 2012; Windschitl 2017; Zembal-Saul et al. 2013). The first round of coding by two co-authors led to additional generations of codes specific to the storybooks in our sample. Subsequent rounds of inter-rater reliability (IRR) coding further refined our coding document. After two rounds of IRR, two coders reached >90% agreement in *investigation questions and planning* and *carrying out investigations* categories. After three rounds of IRR, the same two coders reached 71% agreement on the *evidence-based explanation* codes. Discussion of disagreements improved the coding document and the remaining storybooks were checked by the team to reach 100% agreement on *evidence-based explanations*, improving reliability for this category.

We analysed the data by generating descriptive statistics for categories and codes to look for themes in how students used science practices in their storybooks. We also created a new category, using existing codes, to identify storybooks that presented coherent investigations (Plummer and Tanis Ozcelik, 2015). A coherent investigation included an investigation question, sufficient data, and a claim based on evidence (with or without reasoning). Further, our definitions for each of these codes include the requirements that data are relevant to the investigation question and the claim is in response to the question, resulting in a coherent investigation.

6 Findings

We begin with presenting themes in how students included key science practices (investigation questions, planning and carrying out investigations, constructing explanations) in their storybooks. We conclude with a discussion of students' use of coherent science inquiry investigations. Overall, we found that while the majority of students were able to plan coherent investigations in their storybooks, other students' storybooks were missing key elements of a coherent investigation. However, this improved over the three years we taught the course.

Table 3. Science practices codebook for analysis

Categories	**Codes**	**Definitions**
Investigation Questions	Descriptive question	Characters pose a question about descriptive qualities of a phenomenon (e.g., a "what does it look like" or "how long will it take" question).
	Causal question	Characters pose a question that asks for a causal mechanism for a phenomenon (e.g., "why did that happen").
Planning and carrying out investigations	Sufficient data gathered	Characters gathered a large enough data set to answer their investigation question. Data must be relevant to the investigation question.
	Insufficient data gathered	Characters gathered data, but the dataset was not large enough to support their claims
	Relevant data gathered	Characters gather data that will actually be useful to answer their questions.
	Data gathered, but irrelevant to question	Characters gather data, but data cannot be used to answer investigation question
	Plan is discussed to gather data	Characters explicitly discuss an observation plan before carrying it out. When coding, specify which of the characters is coming up with the plan.
Constructing explanations	Claim not supported by evidence (no reasoning)	Characters present a claim without discussing the evidence or reasoning supporting the claim.
	Claim supported by evidence with no reasoning	Characters have a claim that is supported by their evidence, but do not use scientific concepts to reason through why their evidence supports their claim. This also applies when descriptive questions that do not require reasoning (even if it is provided) are asked.
	Claim based on evidence supported by reasoning	A claim is presented that is supported by evidence. The character also uses scientific ideas and theories to develop a connection between their claim and evidence. This reasoning can be considered an attempt rather than rigorous use of theory/model to support claim and evidence.
	Claim presented with reasoning, but no evidence	A claim is presented without a discussion of the evidence behind it. The character presents scientific ideas and theories to support their claim.
	Claim not coherently matched to evidence or reasoning	A claim is made, but not supported by evidence or reasoning. The evidence and/or reasoning is there, but it does not support the claim (e.g., not aligned to the claim).
Development of reasoning	Main character involved in developing model	The main character develops an explanatory model as part of their explanation (claims, evidence, or reasoning). A model is a sense-making tool that can be used to predict and explain. Can be physical, represented in the story in a character's thoughts, or through character dialogue.
	Main character describes but does not develop model	Main character presents a model of the observed phenomenon but does not provide any indication of where the model came from.
	Model developed by someone else in story	Another character/authority figure develops a model for the main character.

6.1 Topics and question types

Students were instructed to write their storybooks "centered around a Solar System astronomy investigation but [it] does not have to be one of the specific investigations we did in this class" (*Storybook Assignment Sheet*). And while the students often did not follow the exact questions or procedures in their storybooks that the class used when conducting investigations, all topics the students chose were based on investigations we engaged with during the semester. The majority of the students wrote a story investigating the phases of the moon (n=37, 59%). The next most frequent topics were the motion of the Sun, the planets, or the constellations in the Earth's sky (n=11, n=17%), physical properties of the planets (n=6, 9%), craters on solid worlds (n=4, 6%), and light pollution (n=4, 6%). One additional student wrote a story asking how astronomers collect data.

We further analysed whether students posed descriptive questions or causal questions. A descriptive question might seek to describe a pattern in a phenomenon, such as the question in Courtney's storybook: "How long will it take for the Moon to complete its full cycle and go back to its full self?" Nearly a third of the students posed descriptive questions in their storybooks (n=20, 32%). A causal question goes beyond a description to ask about the underlying cause of the observations or pattern in the phenomenon. For example, Mindy's story asks "I wonder why the Moon looks different on different days?" Most students posed causal questions (n=37, 59%). Students were strongly encouraged, through written instructions, in-class discussion about the assignment, and peer review, to write stories that included reasoning in their evidence-based explanations. Providing reasoning suggests that the character investigating the phenomenon is responding to a causal question. Finally, even though students received feedback on a story outline, six students (10%) did not pose an investigation question in their final storybook.

6.2 Planning and carrying out investigations

We also considered how the students provided opportunities for the character to gather data that was both relevant to their investigation question and sufficient for answering that question. The majority of students (n=50, 79%) included relevant data gathered for their investigation. Further, most of the students (n=39, 62%) were also coded as having shown the character gathering sufficient data to answer their investigation question. For example, Diane's storybook "Dolphie and the Stars" follows a dolphin as she answers a question about light pollution, "Why do the stars look different in the sky?" Throughout the story, Dolphie makes observations in several locations including near a large city, near a factory putting out smoke, in a remote area away from cities, and during cloudy weather. Her observations are both relevant and sufficient to support a general claim about why we see different brightnesses or densities of stars in different locations. In contrast, Kaitlyn's storybook, "Max and the Moon," included relevant but insufficient data to answer the investigation question: "How many days are in a Moon's cycle?" The character makes observations of the Moon and records these in his science journal; however, he only observes for 16 days before concluding that the moon's cycle takes one month. The character made an error we frequently observed in both the students' storybooks about the lunar phases and their earlier reports about the lunar phases: they provided only half of a cycle's worth of data but attempted to infer the entire length of the cycle rather than providing sufficient data to support the claim. Only three students' storybooks did not describe a process of data collection. Thus, a majority of the students demonstrated an understanding of the necessity of gathering sufficient, relevant data as important for a scientific investigation—as told through the format of a children's storybook.

6.3 Constructing explanations

Nearly half of the students (n=30, 48%) included a claim supported by both evidence and reasoning developed by one or more characters in the story. For example, in Bethany's story "Sola and the Sun" a young girl, Sola, notices the change in the Sun's location in the sky after painting it at sunrise and sunset. After her father suggests that she investigate her question "How does the Sun appear to move during the day?," Sola uses a compass to track the Sun and paint its location throughout the day (see Figure 2). She uses this evidence to construct a claim: "Look Dad the Sun looks like it moves in a circle! Look how high it was at lunchtime!"

To construct her reasoning, Sola asks her teacher, who tells her to look at a globe to show how the Earth's spinning is "what makes the Sun move in a circle." And while Sola builds on this model for her reasoning, elaborating the explanation for her father, the main character initially relied on outside experts to construct the reasoning. This use of an expert to provide the scientific reasoning in the explanation was relatively common among the preservice teachers (n=27, 43%); fewer storybooks included main characters developing the reasoning for themselves (n=11, 17%). Our goal was for the students to write stories in which the main characters took ownership in developing the reasoning for their explanations as this would provide the reader with a more sophisticated view of modelling practices and scientific reasoning than if an expert were to just tell the main character the reasoning piece as a fait accompli. Therefore, in class, we encouraged students to write stories where the main character collaboratively developed the scientific reasoning, possibly in with a more knowledgeable other or with other characters of similar knowledge level. However, most of our students are likely to have experienced a traditional form of science education where, even if they performed investigations to collect and analyse data, the teacher presented them with the completed scientific model and explained the reasoning for them. In our course, we worked with students, guiding them to develop scientific models they could use to explain their observations. Thus, for many students, our course may be one of the first opportunities to learn science by developing their own scientific models—an important feature of the *NGSS* (McNeill et al., 2017).

An additional 12 students (19%) included a claim supported by evidence without the inclusion of reasoning. For example, in Abby's storybook "Finn's Journey of Finding the Moon," a young turtle, Finn, notices the Moon in the sky appearing to get larger on subsequent nights. He decides to make observations of the Moon to answer his question "How does the appearance of the Moon change?" (descriptive question). Each night, Finn drew pictures of the Moon and began to notice the Moon becoming bigger, until after the 15th night, it started to become smaller.

Figure 2. Image from Bethany's storybook showing a record of her evidence (the location of the Sun throughout the day) to support her claim in the text. She uses the Earth's rotation as reasoning to make sense of how her evidence supports her claim.

On the 31st night, he noticed that the Moon appeared as it did on his first observation. Finally, Finn analyses his observations in order to state a claim: "I realized that in a 30 day cycle, the Moon goes through different phases that cause the appearance to change. After this 30 day cycle, the cycle repeats." Abby shows how Finn used careful observation to figure out the cycle of the phase of the Moon, but concluded her storybook without her character providing reasoning for the explanation.

Students who included a claim supported by evidence without reasoning in their storybooks often posed descriptive questions about the qualities of a phenomenon such as "Are Jupiter and the other planets like our planet, Earth?" (Leah's storybook). It may be that due to the nature of their descriptive investigation questions, their explanations in storybooks did not include reasoning. However, when combined with the students that also included reasoning in their explanations, we found that about two-thirds of our participants (67%) demonstrated how evidence-based claims are important aspects of a scientific investigation through the storybook medium.

6.4 Developing a coherent science inquiry investigation across a storybook

Our primary goal throughout the semester was supporting students' ability to communicate their investigations as coherent narratives, from investigation questions through evidence-based explanations. We identified those storybooks which included CSII: using a story to connect an investigation question, with relevant and sufficient data to answer that question, and constructed a claim based on evidence (with or without reasoning). For example, Ellen wrote a creative story that includes a CSII in her storybook, "The Moon Girl." In her story, a King has a playground which he has forbidden his citizens from entering. The main character, Melissa, plays on his playground, is caught, and is locked in a tower. She will not be let out until she figures out the phases of the Moon and is instructed to "determine how and why the phases of the Moon appear to change or look different over time?" (causal investigation question). To record her data, Melissa drew pictures of the Moon on the wall each day for a month (sufficient data). Over time, she began to notice a pattern in her observations as the Moon began to grow to Full and then decrease back to new (claim based on evidence). Yet, once she figured out the pattern, the King told her she must figure out why this happens, before she can be released. With some guidance from the King's son, Melissa realizes that it is the Sun that illuminates the Moon and that its appearance changes as it orbits the Earth (model-based reasoning). She is released and allowed to play on the playground for the rest of her life.

The majority of students (n=37, 59%) wrote stories featuring CSII, thus enacting how a scientific question can be used to determine what evidence is relevant and sufficient to answer a question, and use that evidence to construct a claim. Further, this percentage improved over time: Year 1, 6 students (35%); Year 2, 14 students (64%); Year 3, 17 students (71%). This suggests that we, the faculty teaching the course, gained insight into ways to further support the students after the first year of the course to better scaffold their understanding of coherent investigations.

7 Conclusions and Recommendations

Our findings demonstrate that, after participating in the "Astronomy for Educators" course, the majority of the preservice teachers were able to enact science practices that led to a coherent, astronomy-based investigation, through the medium of a dual-purpose children's storybook. This supports previous findings that suggest purposefully designed content courses for preservice elementary teachers can play a role in supporting their development towards understanding science as an evidence-based process of explaining phenomena (Avraamidou and Zembal-Saul, 2010; Haefner and Zembal-Saul, 2004). While we were not able to follow these preservice teachers into their first experiences teaching science to elementary students, this provides some evidence that they may enact a use of these science practices in future lesson planning. Prior research indicates that increased domain-specific knowledge influences the level of coherence in inquiry-based lessons produced by future teachers (Plummer and Tanis Ozcelik, 2015). Thus, one useful feature of this course could be how we first supported the preservice teachers' in developing domain-specific knowledge alongside the ways they learned to enact the practices of science. Developing their knowledge may have helped them make sense of coherent astronomy investigations which they represented in curriculum material through dual-purpose children's storybooks.

Yet, a large percentage of students (41%) did not write coherent investigations in their storybooks. Throughout the semester, some students struggled to make connections between asking a scientific question, providing sufficient data, and using this to generate evidence-based claims. This may have continued through into their attempts to write an investigation in storybook form. In addition, some students may have had difficulty translating a new or limited understanding of science practices into a novel format: children's storybooks. However, our students' performance in writing storybooks with coherent investigations improved from 35% in Year 1 to 71% in Year 3. This suggests that over time, we (the faculty teaching the course) improved the ways we engaged our students in coherent scientific investigations, through their participation in a content course designed around a coherent science content storyline. Across each year of the course, instructors supported students through multiple investigations, rounds of sharing in small-group and whole-class conversations, and three reports that used a scientific (rather than creative) format for expressing this same idea of a coherent investigation. Specific improvements made between the Year 1 and 3 included: providing an example report that illustrated how to write about a coherent investigation, increasing how frequently we prompted students to make explicit connections during ongoing class investigations and presentations, providing additional explicit examples of our own of the connections between question, data collection, and explanation throughout the semester, and improving our own understanding and insight into the students' difficulties with engaging in science practices which allowed us to provide more timely support and more helpful feedback. These additional prompts and examples were woven throughout the semester which provided students with additional opportunity to practice doing science with their peers. The social nature of learning in our course contributed to how they were able to translate the enactment of science practices into use in storybook form (Ford, 2008; Vygotsky, 1986).

One of the challenges students had in creating coherent science inquiry investigations was in generating an investigation question for their storybook, even with guidance provided by the instructors. Prior research suggests that the creation of research questions is often the most difficult stage of a scientific investigation for students (Slater et al., 2008). Further, while students in our course were provided with several examples of investigation questions, they did not practice generating investigation questions as part of their in-class investigations. This finding suggests that preservice elementary teachers may need more practice generating investigation questions for coherent inquiry investigation. One limitation was that we did not explicitly teach the students the difference between descriptive and causal questions. More explicit instruction and opportunities to practice generating and using each kind, such as generating a story for both a descriptive and causal investigation, may have helped students understand why each type of investigation is important to science.

Another challenge students had, in generating coherent investigations in their storybooks, was to provide sufficient evidence to support their claim. Most of the problems in students' use of sufficient evidence occurred in storybooks relating to the lunar phases (the most frequent topic) where students included data for half the cycle or less. Students may not have understood how the Moon's cycle was determined from the observations or may have felt that an observer is justified in making a claim that infers the pattern from a limited data. This suggests students may have needed experience with the practice of rebuttal in which students debate which alternative claims best fit the existing evidence (Zembal-Saul et al., 2013). Such an approach could help them appreciate why additional data is needed as evidence to support their claims about the length of the Moon's cycle and to rule out alternative explanations.

And while many students included a claim supported by evidence in their storybooks, fewer provided reasoning that drew on a scientific model or science principle to explain why the phenomena occurred. For some of these students, their choice of descriptive investigation questions may have led to explanations without reasoning. Similar to prior research with middle (McNeill and Krajcik, 2007; McNeill et al., 2006) and high school students (McNeill and Pimentel, 2010), our students found the use of scientific reasoning to be more challenging than providing evidence for their claims. Science courses for preservice teachers may need to support preservice teachers to understand the importance of developing causal questions as a step towards understanding the development of reasoning in evidence-based explanations. In addition, providing opportunities for students to develop hypotheses based on science principles as part of the investigation process may help deepen their understanding of how reasoning is used to construct evidence-based explanations.

Based on our findings, we recommend that other universities offer content courses for preservice teachers that engage them in multiple cycles of coherent inquiry investigation to support their understanding of content and enactment of science practices. We designed this course for preservice teachers to provide multiple opportunities for their engagement in coherent inquiry investigations across the course in astronomical concepts through ongoing classroom collaboration. The scaffolding we offered across the course is consistent with previous studies of astronomy courses that demonstrate how specific guidance significantly supports college students' understandings and science practices (Lyons, 2011; Sibbernsen, 2014; Slater et al., 2008). Our results build on and extend these earlier findings by showing how such a purposefully designed course with multiple cycles of experience to engage in inquiry investigation may lead to a deeper ability to enact science practice, i.e., coherent science inquiry investigations. Explicit in the enactment of these cycles are repeated opportunities to discuss the use of evidence, generation of claims, and application of reasoning among their peers; these practices are key elements to learning in sociocultural theory (Vygotsky, 1986) (Vygotsky, 1998) and reflect how science is learned through the enactment of practices (Ford, 2008).

We also found that a novel assessment format—writing a children's storybook—can be used to assess preservice teachers' ability to enact science practices. Using this assessment, we found evidence of preservice teachers applying knowledge of coherent inquiry investigations to their science stories. This finding provides support to arguments (Murmann and Avraamidou, 2014; Plummer and Cho, 2020; Pringle and Lamme, 2005) about the value of stories as useful learning tools in various learning contexts. We recommend using stories and storybooks as a creative way of assessing student learning. In addition, this novel assessment has the potential to support preservice teachers' future careers in that they can use this form of assessment with their future students.

8 Limitations and Future Research

One limitation of our study is that we did not have an opportunity to measure our students' enactment of science practices at the beginning of the semester in a way comparable to our measurement using the storybook format. This limits the extent to which we can make claims about their growth across the semester. In future iterations of this study, we might consider our research design to capture participants' prior knowledge and abilities in order to track development throughout the course in order to better understand the effects of our intervention in supporting preservice teachers' development of their own curriculum materials. In addition, we are limited in our understanding of what the students learned about astronomy during this course and how their astronomy knowledge influenced their topic choice or their development of science storybooks. We might find that preservice teachers who made greater content gains were more likely to have coherent investigations in their storybooks. This aligns with Plummer and Tanis Ozcelik (2015) who found that preservice teachers who better understood astronomy developed more coherent inquiry investigations. Future research could investigate how preservice teachers' content knowledge may be correlated with the quality of their storybooks.

We also did not include follow-up research on how the preservice teachers use their storybooks as curriculum materials in their teaching to support students' science practices and science inquiry. Avraamidou and Zembal-Saul (2010) illustrate how early-career elementary teachers from the same teacher education program who took different kinds of coursework (typical college lecture-based science courses versus science courses specifically designed for prospective elementary teachers) taught science differently in their first year of teaching. A preservice teacher who took specially designed, inquiry-based science courses in her teacher preparation program was able to modify the curriculum materials for her students and emphasize the discourse of scientific inquiry through investigation. Avvraamidou and Zembal-Saul's study suggests that the inquiry-based preparation our students received may affect choices they make in the beginning of their elementary science teaching careers. Future research could investigate how preservice teachers use curriculum materials they have authored, such as storybooks, to facilitate coherent investigations during field-work experiences. Future research might also consider how experience writing a science inquiry-oriented storybook might shape preservice teacher's choices in curriculum design or storybook choice in their first years of teaching.

In this research, our students were asked to write coherent investigations in story form on the same topics as they had learned in the course. We are limited in our understanding of whether students developed an understanding of the features of a CSII, separate from the particular investigations from our course, or if they were just repeating the same or similar investigations within their fictionalized storybook format without fully understanding why they were including each aspect of their investigation. Bamberger and Davis (2013) point to this issue in their study on how 6th-grade students' ability to transfer modelling performance across content areas. They tested their students' modelling performances and conceptual understanding of three content areas: smell, evaporation, and friction. Through model-based instruction, the researchers taught smell and evaporation but not friction. They found that the students improved their modelling performance in all three content areas. Bamberger and Davis found that middle school students' modelling practices can be transferred to a new content area when learned in the context of science practice-based instruction. Future research could conduct a comparison between writing a coherent investigation storybook about a topic learned in the course and writing a coherent investigation storybook about a science topic familiar to students, but that they did not learn during the course. This could provide insight into the extent to which students extracted the nature of a coherent science inquiry investigation from the context of an astronomy investigation.

Acknowledgements

This study was supported by the National Science Foundation through grant no. 0962792. The findings, opinions, and recommendations expressed in this manuscript are those of the authors and do not necessarily represent the views of the National Science Foundation.

References

Avraamidou, L. and Zembal-Saul, C. (2010). In search of well-started beginning science teachers: Insights from two first-year elementary teachers. *Journal of Research in Science Teaching*, 47(6):661–686. https://doi.org/10.1002/tea.20359.

Bamberger, Y. M. and Davis, E. A. (2013). Middle-school science students' scientific modelling performances across content areas and within a learning progression. *International Journal of Science Education*, 35(2):213–238. https://doi.org/10.1080/09500693.2011.624133.

Brown, M. (2009). Toward a theory of curriculum design and use: Understanding the teacher-tool relationship. In Remillard, J. T., Herbel-Eisenmann, B. A., and Lloyd, G. M., editors, *Mathematics teachers at work: Connecting curriculum materials and classroom instruction*, pages 17–37. New York, Routledge.

Buck, L. B., Bretz, S. L., and Towns, M. H. (2008). Characterizing the level of inquiry in the undergraduate laboratory. *Journal of college science teaching*, 38(1):52–58.

Davis, E. A. and Krajcik, J. S. (2005). Designing educative curriculum materials to promote teacher learning. *Educational researcher*, 34(3):3–14. https://doi.org/10.3102/0013189X034003003.

Donovan, C. A. and Smolkin, L. B. (2002). Considering genre, content, and visual features in the selection of trade books for science instruction. *The Reading Teacher*, 55(6):502–520. https://www.jstor.org/stable/20205092.

Forbes, C. T. (2011). Preservice elementary teachers' adaptation of science curriculum materials for inquiry-based elementary science. *Science Education*, 95(5):927–955. https://doi.org/10.1002/sce.204442.

Ford, M. (2008). Disciplinary authority and accountability in scientific practice and learning. *Science Education*, 92(3):404–423. https://doi.org/10.1002/sce.20263.

Haefner, L. A. and Zembal-Saul, C. (2004). Learning by doing? prospective elementary teachers' developing understandings of scientific inquiry and science teaching and learning. *International Journal of Science Education*, 26(13):1653–1674. https://doi.org/10.1080/0950069042000230709.

Lyons, D. J. (2011). *Impact of Backwards Faded Scaffolding Approach to Inquiry-Based Astronomy Laboratory Experiences on Undergraduate Non-Science Majors' Views of Scientific Inquiry.* PhD thesis, University of Wyoming.

McNeill, K., Berland, L., and Pelletier, P. (2017). Constructing explanations. In Schwarz, C., Passmore, C., and Reiser, B., editors, *Helping Students Make Sense of the World Using Next Generation Science and Engineering Practices*, pages 205–228. NSTA Press: Arlington, VA.

McNeill, K. L. (2011). Elementary students' views of explanation, argumentation, and evidence, and their abilities to construct arguments over the school year. *Journal of Research in Science Teaching*, 48(7):793–823. https://doi.org/10.1002/tea.20430.

McNeill, K. L. and Krajcik, J. (2007). Middle school students' use of appropriate and inappropriate evidence in writing scientific explanations. In Lovett, M. C. and Shah, P., editors, *Thinking with data*, The Proceedings of the 33rd Carnegie Symposium on Cognition, pages 233–265. Mahwah, NJ: Erlbaum.

McNeill, K. L., Lizotte, D. J., Krajcik, J., and Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *The Journal of the Learning Sciences*, 15(2):153–191. https://doi.org/10.1207/s15327809jls1502_1.

McNeill, K. L. and Pimentel, D. S. (2010). Scientific discourse in three urban classrooms: The role of the teacher in engaging high school students in argumentation. *Science Education*, 94(2):203–229. https://doi.org/10.1002/sce.20364.

Murmann, M. and Avraamidou, L. (2014). Animals, emperors, senses: Exploring a story-based learning design in a museum setting. *International Journal of Science Education, Part B*, 4(1):66–91. https://doi.org/10.1080/21548455.2012.760857.

National Research Council (2012). *Framework for K-12 science education*. National Academy Press, Washington, DC.

NGSS Lead States (2013). *Next Generation Science Standards: For states, by states*. The National Academies Press, Washington, DC.

Plummer, J. D. and Cho, K. (2020). Integrating narrative into the design of preschool science programs. In Gresalfi, M. and Horn, I., editors, *The Interdisciplinarity of the Learning Sciences, The 14th International Conference of the Learning Sciences (ICLS)*, International Society of the Learning Sciences, pages 1585–1589, Nashville, TN.

Plummer, J. D. and Tanis Ozcelik, A. (2015). Preservice teachers developing coherent inquiry investigations in elementary astronomy. *Science Education*, 99(5):932–957. https://doi.org/10.1002/sce.21180.

Pringle, R. M. and Lamme, L. L. (2005). Using picture storybooks to support young children's science learning. *Reading Horizons: A Journal of Literacy and Language Arts*, 46(1):2.

Roth, K. and Garnier, H. (2007). What science teaching looks like: An international perspective. *Educational leadership*, 64(4):16–23.

Roth, K. J., Garnier, H. E., Chen, C., Lemmens, M., Schwille, K., and Wickler, N. I. (2011). Videobased lesson analysis: Effective science pd for teacher and student learning. *Journal of Research in Science Teaching*, 48(2):117–148. https://doi.org/10.1002/tea.20408.

Sibbernsen, K. J. (2014). Impact of collaborative groups versus individuals in undergraduate inquiry-based astronomy laboratory learning exercises. *Journal of Astronomy & Earth Sciences Education (JAESE)*, 1(1):41–52. https://doi.org/10.19030/jaese.v1i1.9106.

Slater, S. J., Slater, T. F., and Shaner, A. (2008). Impact of backwards faded scaffolding in an astronomy course for preservice elementary teachers based on inquiry. *Journal of Geoscience Education*, 56(5):408–416. https://doi.org/10.5408/jge_nov2008_slater_408.

Stewart, S. A. (2013). *The design, enactment, and impact of an inquiry-based undergraduate astronomy laboratory learning environment*. PhD thesis, Syracuse University.

Vygotsky, L. S. (1986). *Mind in society: The development of higher mental process*. Harvard University Press.

Windschitl, M. (2017). *Planning and carrying out investigations*, pages 135–158. NSTA Press, Arlington, VA.

Zangori, L. and Forbes, C. T. (2013). Preservice elementary teachers and explanation construction: Knowledge-for-practice and knowledge-in-practice. *Science Education*, 97(2):310–330. https://doi.org/10.1002/sce.21052.

Zembal-Saul, C., McNeill, K. L., and Hershberger, K. (2013). *What's Your Evidence?: Engaging K-5 Children in Constructing Explanations in Science*. Pearson Higher Ed, New York, NY.

ASTRONOMY EDUCATION JOURNAL

Astronomy Education Journal, 2021, Vol. 01, No. 1

doi: 10.32374/AEJ.2021.1.1.003

RESEARCH ARTICLE

Interactive cosmology visualization using the Hubble Ultra Deep Field data in the classroom

Liam J. Nolan[1,*], Mira R. Mechtley[1], Rogier A. Windhorst[1], Karen Knierman[1], Teresa A. Ashcraft[1], Seth H. Cohen[1], Scott Tompkins[1] and Lisa M. Will[2]

[1]School of Earth and Space Exploration, Arizona State University, Tempe, AZ 85287-1404 and [2]San Diego City College, San Diego, CA 92101

[*]ljnolan@asu.edu

Abstract

We have developed a Java-based teaching tool, "Appreciating Hubble at Hyper-speed" ("*AHaH*"), intended for use by students and instructors in beginning astronomy and cosmology courses, which we have made available online. This tool lets the user hypothetically traverse the Hubble Ultra Deep Field (HUDF) in three dimensions at over $\sim 500 \times 10^{12}$ times the speed of light, from redshifts $z = 0$ today to $z = 6$, about 1 Gyr after the Big Bang. Users may also view the Universe in various cosmology configurations and two different geometry modes – standard geometry that includes expansion of the Universe, and a static pseudo-Euclidean geometry for comparison. In this paper we detail the mathematical formulae underlying the functions of this Java application, and provide justification for the use of these particular formulae. These include the manner in which the angular sizes of objects are calculated in various cosmologies, as well as how the application's coordinate system is defined in relativistically expanding cosmologies. We also briefly discuss the methods used to select and prepare the images in the application, the data used to measure the redshifts of the galaxies, and the qualitative implications of the visualization – that is, what exactly users see when they "move" the virtual telescope through the simulation. Finally, we conduct a study of the effectiveness of this teaching tool in the classroom, the results of which show the efficacy of the tool, with over ~90% approval by students, and provide justification for its further use in a classroom setting.

Keywords: Visualization of Relativistic Cosmological Models; Hubble Ultra Deep Field Images; Astronomy Education

1 Introduction

In beginning astronomy courses, many non-science majors appear to have a significant lack of understanding – even after taking introductory courses – of basic concepts such as wavelength, the electromagnetic spectrum, the speed of light, look-back time, redshift, and the expansion of the Universe. We believe this lack of concept acquisition or retention represents a significant shortcoming of currently available teaching tools. While pictures, figures, and other static media are certainly effective at communicating many concepts, they tend to be poor at showing these effects in three dimensions, or those that evolve over time (*e.g.* Sadaghiani 2011). Since virtually all cosmological effects require very large time or distance scales to become apparent, a different teaching medium is needed in this case. This is reflected by the well-established need in astronomy education to support the development of spatial skills, which strongly correlates with performance in all STEM disciplines (Cole et al., 2018).

In addition, the education landscape is changing at a breakneck pace around us, with online learning quickly becoming a preferred option for reasons of convenience, access, and affordability, especially to students of limited means. At its extreme, in the case of large-scale threats to safety, online learning becomes mandatory, as has been so dramatically underlined by the recent outbreak of COVID-19. As millions of students around the world moved from learning in a classroom to learning at home, it became evidently clear that modern classes require new tools for education. Nowhere is this more needed than in the laboratory-type complement to lecture-based learning, and online virtual tools stand to serve this purpose exceptionally well (*e.g.* Hoeling 2012).

"Appreciating Hubble at Hyper-speed" (*AHaH*) is an educational tool that aims to address these issues of concept acquisition and retention by providing a visual and interactive learning medium. The project uses data from the Hubble Space Telescope (HST) Cycle 12 Project "GRAPES" (Grism-ACS (Advanced Camera for Surveys) Program for Extragalactic Science; Pirzkal et al., 2004) to build a redshift-sorted database of over 5000 galaxies within the Hubble Ultra Deep Field (HUDF). As a simplified, brief explanation, if one begins Hubble's Law of the Universal Expansion (as per the FLRW model[1]),

$$D = \frac{v}{H_0} \tag{1}$$

and combine with the Doppler approximation as valid for small redshifts $v \approx cz$, one obtains

$$D = \frac{v}{H_0} \approx \frac{cz}{H_0} \tag{2}$$

and can see a galaxy's distance is proportional to redshift, z, divided by Hubble's constant, H_0[2]. As these galaxies range from redshift $z \approx 0.05$ to $z \approx 6$, their distances span nearly 90% of the history of the Universe (Yan and Windhorst, 2004; Bouwens et al., 2006; Cohen et al., 2006; Windhorst et al., 2011). Since these data represent the deepest optical images of the Universe ever obtained, they are thus uniquely suited to help students understand the effects of the expanding Universe.

It is worth noting that for the *AHaH* Application, we developed a custom-balanced RGB version of the original HUDF image to eliminate bright areas being "burned out" and lacking fine detail (Lupton et al., 2004). This custom version is then displayed as semi-transparent, and images are attached to photometric redshifts measured using HyperZ (Bolzonella et al., 2000) using photometry from the publicly available HST optical and near-infrared images (Thompson et al., 2005), as well as spectro-photometric redshifts from Ryan et al. (2007). A comparison of the original STScI color images and our prepared images is shown in Figure 1, and one can see examples of galaxies processed in this way in Figure 2. We explain this in more detail in § 14 (Appendix B).

1 For further discussion of the differences between the Hubble's redshift-distance and velocity-distance laws, see Harrison (1993).

2 We use $H_0 \approx 68$ km/s/Mpc throughout (Planck Collaboration et al., 2018).

Figure 1. A comparison of three images of HUDF galaxy 7556. The left image is that from the original STScI release, clearly showing the bright, burned-out knots characteristic of the standard logarithmic image stretch. The center image is our prepared image using the arcsinh stretch described by Lupton et al. (2004), as it appears in the *AHaH* application. The right image is our prepared image against an artificially imposed chessboard pattern, showing the included transparency. Note that pixels outside the source are all completely transparent, since they have been removed entirely using the `SourceExtractor` segmentation map.

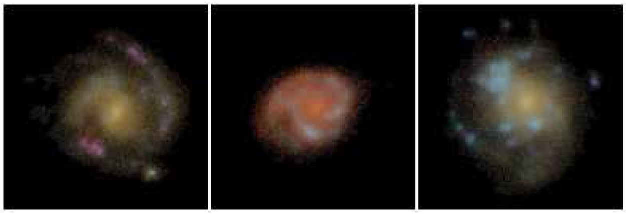

Figure 2. Our prepared images of three galaxies from the HUDF, using the arcsinh stretch described by Lupton et al. (2004). Shown are galaxy 3180 (left), galaxy 5805 (center), and galaxy 6974 (right).

2 Theoretical Formalism

As discussed in detail by Wright (2006), there are a number of different methods for calculating distances in cosmology. For our purposes, the most meaningful of these is the comoving radial distance, D_R, representing the spatial separation of an object and an observer with zero peculiar velocity at a common cosmic time. This distance takes into account the expansion of the Universe, and so is more useful when dealing with distances on very large scales (and thus very large look-back times), as is the case with most galaxies in the HUDF. Henceforth, we shall adopt the convention of referring to the comoving radial distance from Earth to a galaxy as D_R in Gpc[3], and the comoving coordinate distance between two arbitrary points in the coordinate system as r_{ij}.

As we begin to define different cosmological distances, we wish to note that there are many different distance definitions and naming conventions in cosmology which ultimately come down to three distances. For discussions and examples of the variety used in the literature see, for example, Kristian and Sachs (1966), Weinberg (1972), Peebles (1993), Longair (1995), Hogg (1999), Ellis (2007), Etherington (2007), and Ellis (2009). We adopt the names comoving radial distance (D_R), angular size distance (D_A), and luminosity distance (D_L), with D_A and D_L being related by the "distance duality relation." Regardless of terminology or convenient approximations (i.e. Equation (2)), we carefully use the correct distance formalism "under the hood" to produce the visualization in *AHaH*.

Returning to the development of the tool, we also wish to calculate the angular sizes of objects as they would be observed from redshifts other than zero. To do so, we need a formula for the angular size distance, D_A. That is, the distance which satisfies the equation $d \approx \theta D_A$ for an object with transverse linear

3 1 Gpc = 10^9 pc, and 1 pc = 3.26 lightyear.

diameter d subtending an angle θ in the field of view at any redshift in any relativistic cosmology. Following Ribeiro (2005), we note that the correct relation is between solid angle and l^2 over D_A^2, but in the small area of the HUDF we may linearize as previously stated. This is also important to allow the tool to run at a reasonable speed on consumer computers. In a simple Euclidean space, this is the same as the radial distance, but again we must take into account the expansion (and possible curvature) of the Universe, so we must use a separate equation for D_A in the *AHaH* tool. Details on these definitions and equations are given in § 13 (Appendix A).

In addition, we need to consider how we wish to define the coordinate system for the objects within the Java tool. Although we have very deep HST imaging data that allow us to show how the Universe has changed over time, all of these data were collected at a common time (2003/2004–2014). Moreover, the principal distance measure that we have available, the comoving radial distance D_R, also assumes a common cosmic time. Thus the most sensible coordinate system is one with three spatial dimensions that makes all calculations for a common cosmic time, viz. when the data were collected. We can then contract the distances in this "comoving coordinate system" as necessary to simulate observations from redshifts greater than zero. The question remains of how we should derive such coordinates from the data that we have in such a way that they will be useful to us – this is discussed in § 13.3 in Appendix A, prior to deriving the equations. We detail our calculations in full in § 13 (Appendix A), which we include for both completeness and instructional purposes – as these calculations should be comprehensible in an intermediate undergraduate-level cosmology course – but are not necessary to follow for a demonstration of the tool's educational utility. Ultimately, we are able to develop a relationship between angular size distance and redshift, as well as a coordinate system, which is logical with our available data. We use these relationships (with a few simplifications for computational efficiency) to simulate the "motion" of our *AHaH* camera.

3 Standard Display Mode

While some might argue that the equations in § 13 (Appendix A) speak for themselves, we believe it is very instructive to consider the qualitative implications to their user – that is, a description of what exactly we see when we "move" the camera in the Java application. For the sake of completeness, we will also detail a number of cosmological effects that have been omitted from the application due to technical limitations. An example of the standard display mode is shown in Figure 3.

When we move the camera to a certain position in the HUDF (X, Y, Z) data cube, we are in general viewing the Universe as it would appear from that point and at that redshift. We must qualify this statement by noting that the simulation accounts *only* for cosmological effects of changing the camera position – no other dynamical, gravitational lensing, evolutionary, or other cosmological or physical effects are simulated. In this sense, *AHaH* thus truly, though hypothetically, allows the user to travel through the Universe at "hyper-speed." The highest virtual speed of $\sim 500 \times 10^{12}$ times the speed of light that AHaH uses to zoom into the HUDF database allows the observer to travel from z=0 to z $\lesssim$ 6 in a fraction of a minute, rather than in the $\sim$12.9 Gyr needed if the maximum "travel" speed were really limited to the speed of light c, as in the real Universe.

As is covered in many introductory physics and astronomy classes, we use the Euclidean small angle approximation (SAA), where for a fixed object length l and a small angle θ:

$$\theta = tan(\frac{l}{D}) \approx \frac{l}{D}, \qquad (3)$$

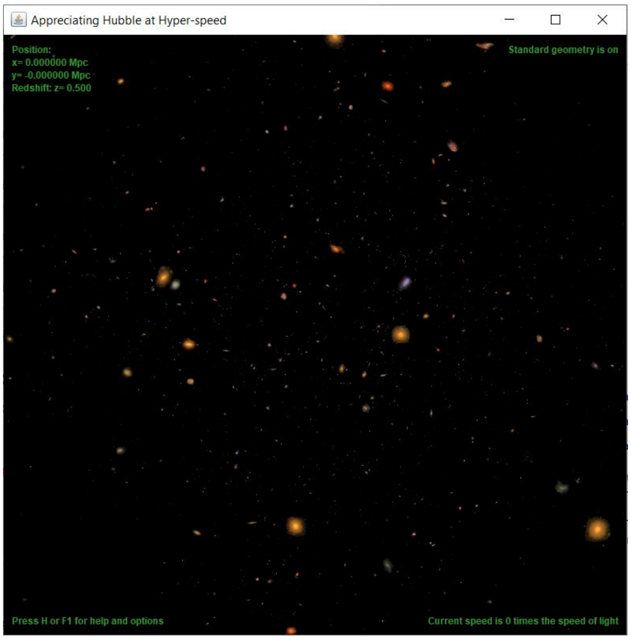

Figure 3. The HUDF data as viewed from redshift z = 0.5 in the *AHaH* application, using standard geometry mode, which properly calculates angular sizes. Note how the image is dominated by luminous red early-type galaxies at moderate redshifts of z $\lesssim$ 1, where the Universe is older than 6 billion years.

where the distance to the object is D. In relativity, we use essentially the same equation, but replace D with D_A, where D_A is a complex function of redshift with a maximum at $z \approx 1.65$. Thus the relativistic SAA is:

$$\theta \approx \frac{l}{D_A}. \qquad (4)$$

The somewhat counter intuitive relationship (called the Θ – z relation) between an object's angular size and its redshift in Relativistic Cosmology is readily apparent in the standard display mode. If a user slowly increases the redshift of the camera, high redshift objects will begin to decrease in angular size and move toward the center of the display, eventually reaching a minimum angular size (at redshift z$\simeq$1.65 in standard ΛCDM cosmology (Planck Collaboration et al., 2018)), and then increasing. Also visible are the effects of galaxy evolution and merging over time. For example, when viewing the Universe from redshift z = 0.5 as in Figure 3, there are many large spiral and elliptical galaxies visible. Note how the Universe is dominated by luminous red, early-type galaxies at moderate redshifts of z $\lesssim$ 1 in Figure 3. However, when viewing the Universe from redshift z = 1.5 as in Figure 4, the *AHaH* screen is dominated by small and compact blue galaxies. When zooming into the data at z $\gtrsim$ 1.5, many of these objects are blue irregular, merging and/or star-forming galaxies. In Figure 4, all red ellipticals of Figure 3 are now "behind us". This Universe at z $\gtrsim$ 1.5 is indeed the actively star-forming Universe, i.e., the first 4 billion years after the Big Bang.

It should be noted that the application does not make calculations for cosmological surface brightness dimming or changes in color due to redshift or spectral evolution. While certainly feasible to simulate, performing such image manipulation techniques on large numbers of galaxies in real-time is currently too difficult for consumer computers. Moreover, we must also recall that the HUDF data are limited in both magnitude and effective horizon by what could be observed from low Earth orbit. When we view the data from redshifts other than zero, we would expect to see more galaxies overall – includ-

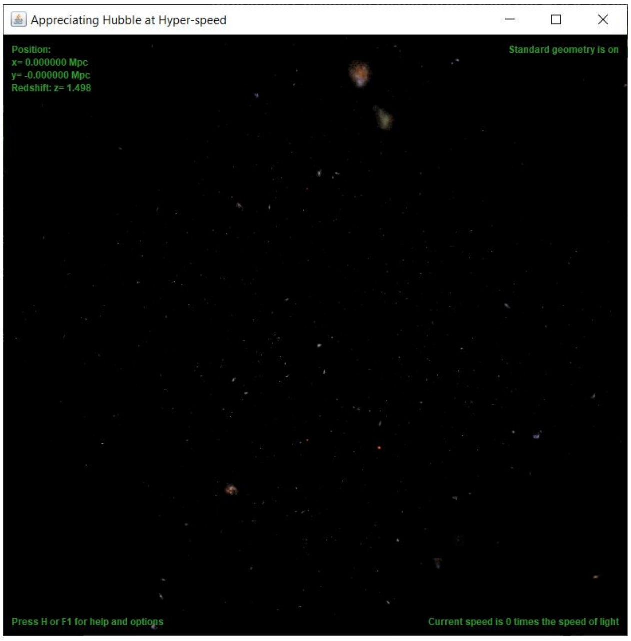

Figure 4. The HUDF data as viewed from redshift $z = 1.5$ in the *AHaH* application, using standard geometry mode. Note how this image is dominated by blue irregular and merging star-forming galaxies, and that all red ellipticals of Figure 3 are now "behind us". This Universe at $z \gtrsim 1.5$ is the actively star-forming Universe, where the Universe is younger than 4 billion years.

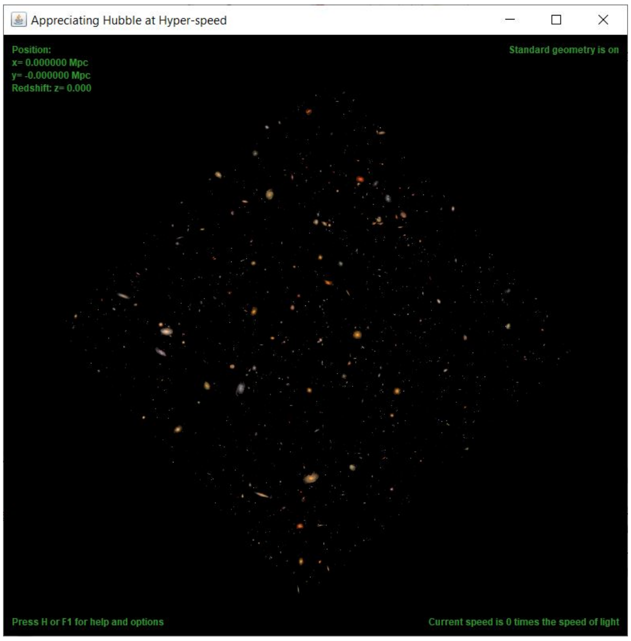

Figure 5. The HUDF data as viewed from redshift $z = 0$ in the *AHaH* application, using standard geometry mode. This displays the entire HUDF, for comparison with Figures 6 and 7.

ing fainter galaxies – than are present in the current HUDF data. We could choose to simulate these objects as extensions of our data set if desired, but we felt this would not be particularly instructive, and could lead to potential confusion, as fainter galaxies would have to be continuously simulated in increasing numbers by the computer below Hubble's detection limit when traveling from z=0 today to $z \lesssim 6$ (~1 Gyr after the Big Bang). Moreover, such simulations have a high degree of uncertainty, and by significantly increasing the size of the data set, would add prohibitively to the computation times. Likewise, we have chosen not to simulate galaxies outside of the original field, which would of course enter the camera's field of view as the user pans around.

4 Static Geometry Mode

When a user presses the "G" key in the *AHaH* Java tool, they are told that they are viewing the simulation with "unexpanding angular geometry" turned on. What this means specifically is that angular sizes as derived above are no longer affected by the scale factor or curvature of the Universe. After we develop our original coordinates, as in Equation (14), all calculations for angles are simply done with $\theta = \theta_E$. This has the visual effect of all galaxies appearing smaller and closer to the center of the viewport (as can be seen in Figure 6 as compared to 5, since all initial angles have been contracted by an expansion factor of $(1+z)$ (when the curvature energy density Ω_K is zero; see § 5 for an explanation of the main energy densities at play in Relativistic Cosmology). In this static case, galaxies will also simply increase in angular size as we approach them, as opposed to the angular sizes of high-redshift objects in the real ΛCDM Universe (Planck Collaboration et al., 2018), which decrease, reach a minimum, and then *increase* again in angular size as the camera's redshift *increases*.

This static mode of viewing the simulation has no physical analogue – it is simply meant to convey to the user that there are non-Euclidean aspects of the Universe's geometry, and that the

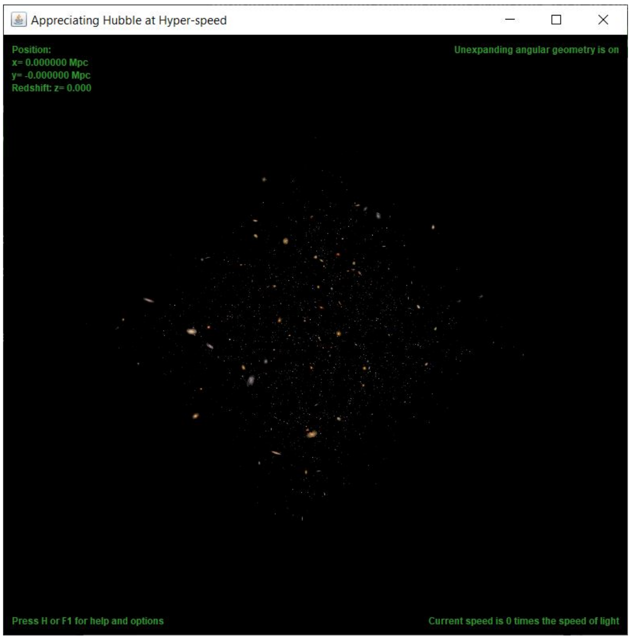

Figure 6. The HUDF data as viewed from redshift $z = 0$ in the *AHaH* application, using static geometry mode. This displays the entire HUDF, and one can see the evident "contraction" of the field of galaxies due to the lack of non-Euclidean geometry in the real HUDF data. (That is, the HUDF data in a relativistically expanding cosmology – when shown in Euclidean geometry – compresses all objects towards the image center, since the square cylindrical volume is now not undergoing the expansion, as it should).

angular sizes that we observe in the present have been made larger due to the universal expansion. One should note that this display mode only considers expansion as it relates to angular size – the comoving radial distance is still calculated using the redshift and curvature factors that would not be present in a strictly Euclidean Universe. That is, in the static display mode, we assume that the Hubble Law distance, $D = v/H_0 \approx (c/H_0)z$, is simply a Euclidean distance unrelated to expansion. This is primarily because our method of calculating the comoving radial distance relies using all object redshifts, which is a phenomenon specific only to an expanding Universe, and is therefore the only way we could calculate the distances for all galaxies in other – hypothetical – Universes.

5 Exploration of Extreme Cosmologies

One additional capacity of the *AHaH* software we wish to note is the representation of wildly different universes from our own, in terms of the cosmological parameters used in the calculations above. As we have already noted, many students have difficulty with relatively esoteric concepts of the energy density parameters of the Universe, and how different portions dominate cosmic behavior over time. To briefly summarize for the reader, Ω_M, Ω_R, Ω_Λ, and Ω_K are the main cosmological parameters, which are the fractions of the Universe's total average energy density that are attributable to matter (M), radiation (R), dark energy (Λ), and the curvature of the spatial geometry (K), respectively. It is assumed these are the only relevant contributions to the total energy density Ω_{Tot}, i.e. $\Omega_{Tot} = \Omega_M + \Omega_\Lambda + \Omega_R + \Omega_K$ (with a spatially flat Universe having $\Omega_{Tot} \equiv 1$ with $\Omega_K = 0$). However, instead of leaving this to verbal and written descriptions (which can be obtuse), relying on a student's ability to mentally visualize these concepts, *AHaH* allows the student to form an approximate image of any cosmology with whatever parameters one could desire. By varying a single parameter, one can see the dramatic (or sometimes less so) impact on our vision of the Universe. We present as an example a massive increase in the vacuum energy density (or Einstein's cosmological constant) Ω_Λ from 0.763 to 2.0 (as in Figure 7), which of course results in a Universe rapidly pushed outwards until just a few galaxies hypothetically remain in Hubble's field-of-view. This value was selected because at noticeably higher values of Ω_Λ, *all* galaxies disappear nearly completely from view sooner. Of course, just as with the static viewing mode described above, this change in parameters does not produce a perfect representation of a Universe with these parameters, and what is simulated by the software has no physical analogue. However, the changes in display under extreme values of different constants is still instructive as to the effects of these different constants on the structure of the Universe.

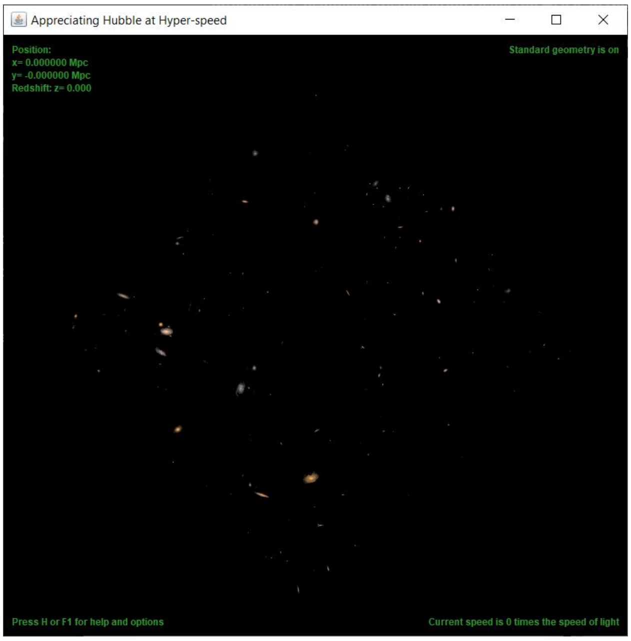

Figure 7. The HUDF data as viewed from redshift $z = 0$ in the *AHaH* application, using standard geometry mode, with a vacuum energy density parameter Ω_Λ value increased from the default value of 0.763 to 2.0. One can see that the extreme value causes most galaxies that are visible in the standard ΛCDM cosmology (Planck Collaboration et al., 2018) to disappear due to vastly increased expansion in this case. This value was selected because noticeably higher values of Ω_Λ cause all galaxies to disappear entirely – we observers likely live in a relativistic Universe with a "little Lambda."

6 Integration in the Classroom

In order to test the effectiveness of this software in a real education environment, we conducted a study in astronomy classes at Arizona State University in the northern hemisphere fall semester of 2020. It should be noted that due to the outbreak of COVID-19, ASU classes were moved into a partially online Synchronous format for part/all of the northern hemisphere fall 2020 semester, thus introducing an additional variable as compared to a typical semester. This pandemic underlines the necessity of developing online laboratory tools which complement virtual instruction, thus we felt continuing with our study was particularly appropriate.

6.1 Methods

We selected the introductory astronomy labs for non-majors (AST 113 northern hemisphere Fall 2020) and for astronomy majors (SES 123 northern hemisphere Fall 2020) for our study in order to focus on the target population mentioned above – students first being introduced to this type of astronomy content. Both classes are typically composed of first-year students, so the level of preparation is similar, though students pursuing astronomy majors tend to come into college with more background information in our experience. We modified the curriculum of all sections of the lab in two different orders, and compared our results between the two orderings of labs. We requested the ASU Institutional Review Board (IRB) for approval to make this study in the AST 113 and SES 123 Lab classrooms, and their approval, containing all conditions for the study, was filed with ASU. Students who enrolled in these modified lab sections were informed that the new programs with our virtual tool would be available, and that this would not affect the course expectations or standards.

We selected several exercises in the typical class curriculum in both labs for replacement with *AHaH* materials based on the advice of the course instructor as to which exercises could have their essential materials folded into other activities or the class lecture component. We then introduced in replacement of these exercises a series of exercises designed by members of our team for *AHaH*, all of which are available online[4]. Students then conducted these exercises either on AST 113/SES 123 Lab or their own home computers, with the same resources available to them as other activities in class (i.e. technical and subject-matter support from Teaching Assistants, instructor, etc.). Reasonable accommodations for learning and access difficulties as pursuant to University policy were made for, *e.g.*, those students who needed to take this course remotely via the Zoom teleconference platform given their presence elsewhere during the COVID-19 pandemic. In addition, for visually impaired students, we made available the Astronomy Sound of the Month webpage for January 2018 [5], which has a full display of the HUDF image and plays a tone when one moves the mouse over a given galaxy, of varying pitch depending on the distance to that galaxy. All lab sections were led for these activities by a combination of Lab Teaching Assistants and the class

4 http://ahah.asu.edu/exercises.html
5 https://astrosom.com/Jan2018.php

instructor, as appropriate.

In some cases, material covered by these *AHaH* exercises is not part of the subject matter typically evaluated by the course. However, the course instructor believed there was sufficient relation to course content, and students were made aware of this slight alteration. These subjects included, *e.g.*: galaxy morphology, Hubble's Law, galaxy evolution, the $\Theta - z$ relation, and cosmological parameters.

6.2 Educational Materials

The first of the two student activities (referred to as "labs") we led students through in this class was on the subject of galaxy morphology[6]. Students use the *AHaH* tool to "fly" through the universe and inspect the morphology of galaxies, either from an instructor-provided list or whichever they find interesting. In the latter case, students are encouraged to pick a variety of galactic morphologies, as they are then asked to classify the galaxies by the Hubble Classification Scheme, and identify various other features. Students then learn more about how different features can lead to conclusions about the populations of stars making up the galaxy. The general learning outcomes produced are:

- Identify galaxies by appearance
- Apply observed colors to the properties of its stars, and
- Understand the differences between the different Hubble classification types.

The second lab was on the subject of "Hubble's Law"[7], by which we mean the linear relationship found by Hubble between distance and redshift at low redshifts. In a similar fashion to the galaxy morphology lab, students move through the HUDF to inspect galaxies either from a list or of their choice, but in this case they use the built-in information pop-up on the galaxy to learn its redshift and comoving radial distance, D_R. Students plot their data points, and then use the "Hubble Law" (the combination of the FLRW model Hubble Law and the Doppler approximation as given in Equation (2)) to calculate H_0 from the slope of their graphs. A spreadsheet which would work out distances for higher redshifts was made available to students who wished to look at such galaxies. The general learning outcomes produced are:

- Understanding the concept of redshift
- Obtaining and performing calculations with data to produce universal properties (such as H_0), and
- Gaining familiarity with the expansion of the universe.

These goals in particular address an established stumbling point in introductory undergraduate astronomy education - the curvature and behavior of the Universe - as described by Coble et al. (2018).

The goals of these labs are well-described by the Anatomy of Disciplinary Discernment as laid out by Eriksson et al. (2014). For many students, the labs will serve to bring them from the first level, Disciplinary Identification (as students are expected to have been previously informed of terms such as "galaxy" and "redshift"), to the second level, Disciplinary Explanation, and be able to show how galactic properties distribute with redshift. For other students the labs may be able to bring them to the third level, Disciplinary Appreciation, as they grasp the "power" of Hubble's Law and other astronomical relations for astronomers to be able to describe the universe itself from the

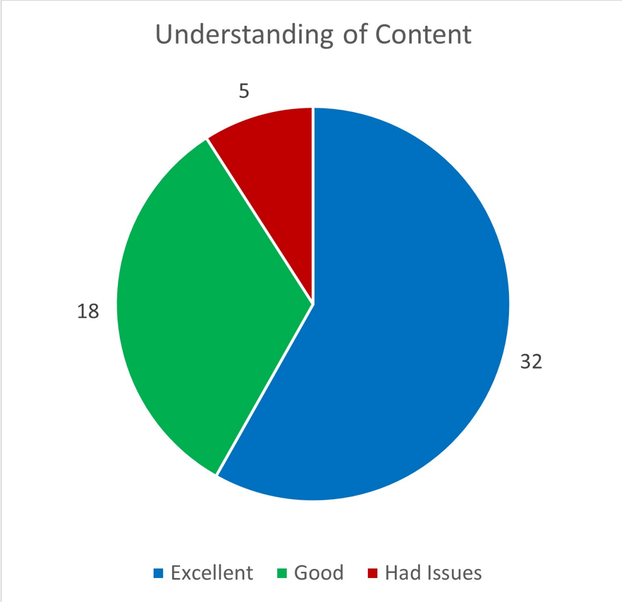

Figure 8. Summary of results from northern hemisphere fall 2020 study, detailed in Table 2, specifically for question on how well the tool helped the students' understanding of the lab content.

trends of galaxies. We also see a few students actually take it upon themselves to learn more about how the standard calculations of the lab are approximations, and how these diverge at high redshifts - thus approaching the fourth and final level, Disciplinary Evaluation. We would also note that a third lab is available on our website concerning galaxy evolution[8], but was not used in our study.

6.3 Analysis and Results

At the end of the northern hemisphere fall 2020 semester class described above, we conducted a survey of students in the enhanced labs for their assessment of the utility of the *AHaH* tool, with questions shown in Table 1, and complete results shown in Figures 8, 9, and 10, and Table 2 for the northern hemisphere fall 2020 class. As can be seen, 90% of students surveyed thought that *AHaH* was good to excellent in helping understanding of lab materials, 92% thought it was good to excellent in making the lab more interesting, and 91% thought that it was good to excellent in answering questions they had on the subject matter. Finally, throughout the semester students consistently expressed to instructors a preference for activities with *AHaH*, as well as greater ease of understanding with the tool. Teaching Assistants regularly reported notably higher amounts of positive feedback on AHaH-related labs as compared to standard lab exercises.

7 Other Program Uses

In addition to in the classroom, *AHaH* has been and continues to be useful in a wide range of applications in STEM education. We have conducted several pilot programs to use this tool in outreach efforts through several events held on and about ASU campus. As in education, while static images have utility in outreach efforts with the general public, people young and old are far more likely to become and stay engaged with dynamic media, such as simulation. As previous studies have shown (*e.g.*

6 http://ahah.asu.edu/exercises/galmorph.pdf
7 http://ahah.asu.edu/exercises/hubble.pdf
8 http://ahah.asu.edu/exercises/galevol.pdf

Table 1. Questions from survey used in northern hemisphere Fall 2020 AST 113/SES 123 program.

Question	Response Type
I understand that my responses will be used – without my name and in aggregate – in a journal article, and I consent to this use.	Acknowledgement
How well did you feel AHaH and related labs aided your understanding of the content covered by these labs?	Multiple Choice on scale of 1 (Poor) to 5 (Excellent)
How well did you feel AHaH and related labs helped in making labs more interesting?	Multiple Choice on scale of 1 (Poor) to 5 (Excellent)
How well did you feel AHaH and related labs helped answer your questions about this material?	Multiple Choice on scale of 1 (Poor) to 5 (Excellent)
Other comments?	Free response

Table 2. Summary of results from northern hemisphere fall 2020 program, with number and percentage of responses to questions, and summary of comments from students. Each of the two categories above were rated on a scale of 1-5, with 1 being not at all and 5 being excellently useful in this regard. We consider here responses of 5 to be "Excellent," 3-4 to be "Good," and 1-2 to be "Had Issues."

Category	Excellent	Good	Had Issues
Understanding of Lab Materials	32 (58%)	18 (32%)	5 (9%)
Interesting for Lab Exercise	31 (56%)	20 (36%)	4 (7%)
Answered Questions	31 (56%)	19 (35%)	5 (9%)
Comments	Excited to use tool as visualization - many students were interested in the differences between galaxies visually, as well as getting an idea of differences in distances.	Interface took getting used to - some students expressed an interest in the tool being redesigned as an app, or with more immediate information pop-ups.	Common difficulties were with installation of the tool.

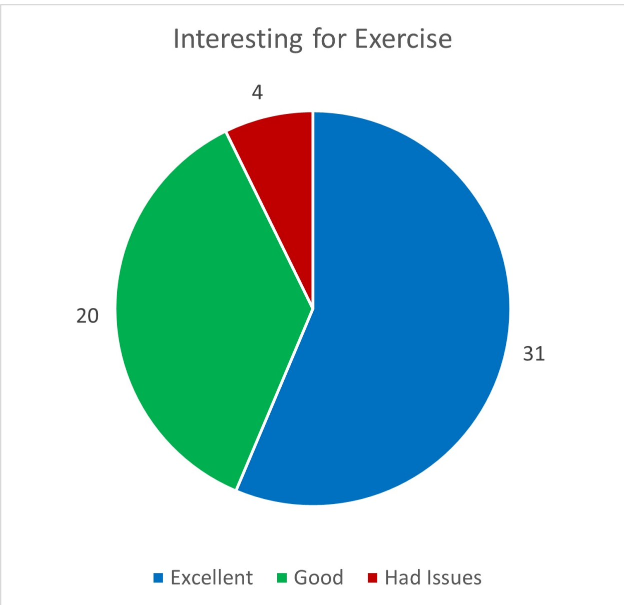

Figure 9. Summary of results from northern hemisphere fall 2020 study, detailed in Table 2, specifically for question on how interesting students found the tool.

Figure 10. Summary of results from northern hemisphere fall 2020 study, detailed in Table 2, specifically for question on how well the tool answered students' questions.

Holzinger et al. 2008), simple dynamic media can improve the acquisition and retention of information. In the case of interactive simulation like *AHaH*, this is likely because the user becomes directly involved in shaping the course of their experience. In our program, this is exemplified by members of the public being able to pick out the galaxies they “fly" towards using *AHaH*, and can learn more details about these. These programs have shown a qualitative increase in participation by the general public in outreach activities, especially among young children or young adults. As many public educators will attest, half the battle is often getting the public to start using an education opportunity. Thus the use of *AHaH* in the classroom is worth consideration by teachers, outreach developers, and organizers.

8 Conclusion

We believe that our *AHaH* software provides students and instructors with an unique ability to interactively visualize many of the effects of a relativistically expanding Universe, among its other capabilities. The application should help clarify these concepts, and allow students to develop a deeper intuitive understanding of the material. Certain cosmological effects – such as bandpass shifting, k-correction, surface brightness dimming, gravitational lensing, and the effects of the magnitude limit and object sizes on the sample completeness limit – have been largely omitted due to computational limitations, but we believe these to be not essential for the understanding of the included effects. For a discussion of these more technical effects, see *e.g.*, Windhorst et al. (2018). In addition, our brief study of the utility of the program in the classroom meets our expectations of its impact on learning, and lends support to our recommendation that further virtual tools be developed in support of online classrooms. We also recommend the use of this tool in other public education and science outreach efforts for its utility in quickly engaging the public.

For the convenience of those who wish to see or modify the particular implementation of the above formulae within the Java software, we have provided source code with the standard distribution of the tool. It is included in the src/ directory of ahah.jar, and may be extracted using the java `jar` utility or any zlib-compatible de-compressor such as `unzip`. The tool may be downloaded from the *AHaH* website[9]. Further details on *AHaH* download instructions and installation are given in the *AHaH* user manual, available on the tool website.

9 Availability of source code and requirements

- Project name: Appreciating Hubble At Hyperspeed
- Project home page: http://ahah.asu.edu
- Operating system(s): Microsoft Windows, Mac OS, or *nix OS
- Programming language: Java
- Other requirements: Java 1.4.3 or higher, 1 GHz processor, 256 MB RAM, Mouse and Keyboard
- License: BSD-like.

Virtually any modification or redistribution of the application is permitted, with the following caveats:

- Any redistribution must retain the original copyright notice and license file, either with the source code or with the documentation in the case of binary distributions.
- The names of the copyright holders, contributors, and associated institutions may not be used to endorse or promote any derivative works without prior permission.

The application's source code is provided with the standard distribution.

10 Availability of supporting data and materials

The data used to develop the teaching tool is available through the sources we cite throughout the paper. As part of our Institutional Review Board agreement, our data may only be published and discussed in aggregate, so we cannot make available our raw survey data.

11 Declarations

11.1 List of abbreviations

- ACS : Advanced Camera for Surveys
- *AHaH* : Appreciating Hubble at Hyperspeed
- ASU : Arizona State University
- CTIO : Cerro Tololo Inter-American Observatory
- GRAPES : Grism-ACS Program for Extragalactic Science
- HUDF : Hubble Ultra Deep Field
- HST : Hubble Space Telescope
- IRB : Institutional Review Board
- ISAAC : Infrared Spectrometer And Array Camera
- NICMOS : Near Infrared Camera and Multi-Object Spectrometer
- VLT : Very Large Telescope

11.2 Ethical Approval

As noted above, we filed for and received approval with the ASU Institutional Review Board to conduct our survey in the AST 113 and SES 123 Lab classrooms, and their approval, containing all conditions for the study, was filed with ASU. Students were informed of the new content that would be part of the class, and that there would be a non-graded, optional survey at the end of class concerning the *AHaH* tool. All participants in the survey, as indicated in our procedures, gave consent that they were over 18 years of age and that their responses would be reported in aggregate.

11.3 Consent for publication

Not applicable, see above.

11.4 Competing Interests

The authors declare that they have no competing interests.

11.5 Funding

We acknowledge student support from the Arizona State University NASA Space Grant (to LJN and MRM). We acknowledge support from Hubble Space Telescope grants HST-GO-10530.07-A, HST-GO-13779.005-A, HST-EO-10530.25-A and HST-EO-13241.001-A from STScI, which is operated by AURA for NASA under contract NAS 5-26555. RAW acknowledges support from NASA JWST Interdisciplinary Scientist grants NAG-12460, NNX14AN10G and 80NSSC18K0200 from GSFC.

9 http://ahah.asu.edu

11.6 Author's Contributions

LN was the general project lead, developed the survey and other in-classroom components, completed IRB procedures, and wrote the majority of the text on background and program applications. MM led development of the *AHaH* code and wrote much of the initial technical paper on the tool's function. RW was in charge of general project management and oversight, as well as geometric formalism, development of related labs, and assisted with paper modification. KK, LW, TA, and RW were in charge of various labs and accommodating the *AHaH* survey. SC helped with the data preparation for the tool, and assisted in paper development. ST gave general project assistance.

12 Acknowledgements

We thank Dr. Ned Wright for helpful discussion early in the project. We sincerely appreciate comments from both reviewers of this paper, which helped us better anchor our work in the historical context, and better suit the audience of AEJ and current educational research. We also thank the wonderful Teaching Assistants of AST 113 and SES 123: Angelica Berner, Katherine Elder, Ebraheem Farag, Connor Gelb, Jake Hanson, Isabela Huckabee, Darby Kramer, Kelley Liebst, and Kyle Massingill. Much of our education design was based in part on the work by Hasper et al. (2015).

References

Beckwith, S. V. W., Stiavelli, M., Koekemoer, A. M., Caldwell, J. A. R., Ferguson, H. C., Hook, R., Lucas, R. A., Bergeron, L. E., Corbin, M., Jogee, S., Panagia, N., Robberto, M., Royle, P., Somerville, R. S., and Sosey, M. (2006). The Hubble Ultra Deep Field. *AJ*, 132(5):1729–1755.

Bertin, E. and Arnouts, S. (1996). SExtractor: Software for source extraction. *AAPS*, 117:393–404.

Bertin, E., Mellier, Y., Radovich, M., Missonnier, G., Didelon, P., and Morin, B. (2002). The TERAPIX Pipeline. In Bohlender, D. A., Durand, D., and Handley, T. H., editors, *Astronomical Data Analysis Software and Systems XI*, volume 281 of *Astronomical Society of the Pacific Conference Series*, page 228.

Bolzonella, M., Miralles, J. M., and Pelló, R. (2000). Photometric redshifts based on standard SED fitting procedures. *AAP*, 363:476–492.

Bouwens, R. J., Illingworth, G. D., Blakeslee, J. P., and Franx, M. (2006). Galaxies at z ∼6: The UV Luminosity Function and Luminosity Density from 506 HUDF, HUDF Parallel ACS Field, and GOODS i-Dropouts. *APJ*, 653(1):53–85.

Coble, K., Conlon, M., and Bailey, J. M. (2018). Investigating undergraduate students' ideas about the curvature of the universe. *Phys. Rev. Phys. Educ. Res.*, 14:010144.

Cohen, S. H., Ryan, R. E., J., Straughn, A. N., Hathi, N. P., Windhorst, R. A., Koekemoer, A. M., Pirzkal, N., Xu, C., Mobasher, B., Malhotra, S., Strolger, L. G., and Rhoads, J. E. (2006). Clues to Active Galactic Nucleus Growth from Optically Variable Objects in the Hubble Ultra Deep Field. *APJ*, 639(2):731–739.

Cole, M., Cohen, C., Wilhelm, J., and Lindell, R. (2018). Spatial thinking in astronomy education research. *Phys. Rev. Phys. Educ. Res.*, 14:010139.

Dahlen, T., Mobasher, B., Dickinson, M., Ferguson, H. C., Giavalisco, M., Kretchmer, C., and Ravindranath, S. (2007). Evolution of the Luminosity Function, Star Formation Rate, Morphology, and Size of Star-forming Galaxies Selected at Rest-Frame 1500 and 2800 Å. *APJ*, 654(1):172–185.

Ellis, G. F. R. (2007). On the definition of distance in general relativity: I. M. H. Etherington (Philosophical Magazine ser. 7, vol. 15, 761 (1933)). *General Relativity and Gravitation*, 39(7):1047–1052.

Ellis, G. F. R. (2009). Re-publication of: Relativistic cosmology (1971), In: General Relativity and Cosmology, Proc. of the International School of Physics "Enrico Fermi," R.K. Sachs (ed), (New York: Academic Press). *General Relativity and Gravitation*, 41(3):581–660.

Eriksson, U., Linder, C., Airey, J., and Redfors, A. (2014). Introducing the anatomy of disciplinary discernment: an example from astronomy. *European Journal of Science and Mathematics Education*, 2:167–182.

Etherington, I. M. H. (2007). Re-publication of: LX. On the definition of distance in general relativity (1933), Phil. Mag., 15, 761. *General Relativity and Gravitation*, 39(7):1055–1067.

Harrison, E. (1993). The Redshift-Distance and Velocity-Distance Laws. *APJ*, 403:28.

Hasper, E., Windhorst, R., Hedgpeth, T., Van Tuyl, L., Gonzales, A., Martinez, B., Yu, H., Farkas, Z., and Baluch, D. (2015). Methods for Creating and Evaluating 3D Tactile Images to Teach STEM Courses to the Visually Impaired. *J. College Sc. Teaching*, 44.

Hoeling, B. (2012). Interactive online optics modules for the college physics course. *American Journal of Physics*, 80:334.

Hogg, D. (1999). Distance measures in cosmology.

Holzinger, A., Kickmeier-Rust, M., and Albert, D. (2008). Model-based gaussian and non-gaussian clustering. *Educational Technology & Society*, 11:279–290.

Kristian, J. and Sachs, R. K. (1966). Observations in Cosmology. *APJ*, 143:379.

Longair, M. (1995). The deep universe. *Saas-Fee Advanced Course23*, page 317.

Longair, M. (1998). *Galaxy formation*. New York : Springer.

Lupton, R., Blanton, M. R., Fekete, G., Hogg, D. W., O'Mullane, W., Szalay, A., and Wherry, N. (2004). Preparing Red-Green-Blue Images from CCD Data. *PASP*, 116(816):133–137.

Peebles, P. J. E. (1993). *Principles of physical cosmology*. Princeton University Press.

Pirzkal, N., Malhotra, S., Ryan, R. E., Rothberg, B., Grogin, N., Finkelstein, S. L., Koekemoer, A. M., Rhoads, J., Larson, R. L., Christensen, L., Cimatti, A., Ferreras, I., Gardner, J. P., Gronwall, C., Hathi, N. P., Hibon, P., Joshi, B., Kuntschner, H., Meurer, G. R., O'Connell, R. W., Oestlin, G., Pasquali, A., Pharo, J., Straughn, A. N., Walsh, J. R., Watson, D., Windhorst, R. A., Zakamska, N. L., and Zirm, A. (2017). FIGS—Faint Infrared Grism Survey: Description and Data Reduction. *APJ*, 846(1):84.

Pirzkal, N., Xu, C., Malhotra, S., Rhoads, J. E., Koekemoer, A. M., Moustakas, L. A., Walsh, J. R., Windhorst, R. A., Daddi, E., Cimatti, A., Ferguson, H. C., Gardner, J. P., Gronwall, C., Haiman, Z., Kümmel, M., Panagia, N., Pasquali, A., Stiavelli, M., di Serego Alighieri, S., Tsvetanov, Z., Vernet, J., and Yan, H. (2004). GRAPES, Grism Spectroscopy of the Hubble Ultra Deep Field: Description and Data Reduction. *APJS*, 154(2):501–508.

Planck Collaboration, Aghanim, N., Akrami, Y., Ashdown, M., Aumont, J., Baccigalupi, C., Ballardini, M., Banday, A. J., Barreiro, R. B., Bartolo, N., Basak, S., Battye, R., Benabed, K., Bernard, J. P., Bersanelli, M., Bielewicz, P., Bond, J. R., Borrill, J., Bouchet, F. R., Burigana, C., Calabrese, E., Carron, J., Chiang, H. C., Comis, B., Contreras, D., Crill, B. P., Curto, A., Cuttaia, F., de Bernardis, P., de Rosa, A., de Zotti, G., Delabrouille, J., Di Valentino, E., Dickinson, C., Diego, J. M., Doré, O., Ducout, A., Dupac, X., Elsner, F., Enßlin, T. A., Eriksen, H. K., Falgarone, E., Fantaye, Y., Finelli, F., Forastieri, F., Frailis, M., Fraisse, A. A., Franceschi, E., Frolov, A., Galeotta, S., Galli, S., Ganga, K., Gerbino, M., Górski, K. M., Gruppuso, A., Gudmundsson, J. E., Handley, W., Hansen, F. K., Herranz, D., Hivon, E., Huang, Z., Jaffe, A. H., Keihänen, E., Keskitalo, R., Kiiveri, K., Kim, J., Kisner, T. S., Krachmalnicoff, N., Kunz, M., Kurki-Suonio, H., Lamarre, J. M., Lasenby, A., Lattanzi, M., Lawrence, C. R., Le Jeune, M., Levrier, F., Liguori, M., Lilje, P. B., Lindholm, V., López-Caniego, M., Lubin, P. M., Ma, Y. Z., Macías-Pérez, J. F., Maggio, G., Maino, D., Mandolesi, N., Mangilli, A., Martin, P. G., Martínez-González, E., Matarrese, S., Mauri, N., McEwen, J. D., Melchiorri, A., Mennella, A., Migliaccio, M., Miville-Deschênes, M. A., Molinari, D., Moneti, A., Montier, L., Morgante, G., Natoli, P., Oxborrow, C. A., Pagano, L., Paoletti, D., Partridge, B., Perdereau, O., Perotto, L., Pettorino, V., Piacentini, F., Plaszczynski, S., Polastri, L., Polenta, G., Rachen, J. P., Racine, B., Reinecke, M., Remazeilles, M., Renzi, A., Rocha, G., Roudier, G., Ruiz-Granados, B., Sandri, M., Savelainen, M., Scott, D., Sirignano, C., Sirri, G., Spencer, L. D., Stanco, L., Sunyaev, R., Tauber, J. A., Tavagnacco, D., Tenti, M., Toffolatti, L., Tomasi, M., Tristram, M., Trombetti, T., Valiviita, J., Van Tent, F., Vielva, P., Villa, F., Vittorio, N., Wandelt, B. D., Wehus, I. K., Zacchei, A., and Zonca, A. (2018). Planck intermediate results. LIII. Detection of velocity dispersion from the kinetic Sunyaev-Zeldovich effect. *AAP*, 617:A48.

Retzlaff, J., Rosati, P., Dickinson, M., Vandame, B., Rité, C., Nonino, M., Cesarsky, C., and GOODS Team (2010). The Great Observatories Origins Deep Survey. VLT/ISAAC near-infrared imaging of the GOODS-South field. *AAP*, 511:A50.

Ribeiro, M. B. (2005). Cosmological distances and fractal statistics of galaxy distribution. *A&A*, 429(1):65–74.

Ryan, R. E., J., Hathi, N. P., Cohen, S. H., Malhotra, S., Rhoads, J., Windhorst, R. A., Budavári, T., Pirzkal, N., Xu, C., Panagia, N., Moustakas, L. A., di Serego Alighieri, S., and Yan, H. (2007). The Galaxy Luminosity Function at z $\sim$= 1 in the HUDF: Probing the Dwarf Population. *APJ*, 668(2):839–845.

Ryden, B. (2017). *Introduction to Cosmology, 2nd Edition*. Cambridge University Press, Cambridge, United Kingdom.

Sadaghiani, H. R. (2011). Using multimedia learning modules in a hybrid-online course in electricity and magnetism. *Phys. Rev. ST Phys. Educ. Res.*, 7:010102.

Thompson, R. I., Illingworth, G., Bouwens, R., Dickinson, M., Eisenstein, D., Fan, X., Franx, M., Riess, A., Rieke, M. J., Schneider, G., Stobie, E., Toft, S., and van Dokkum, P. (2005). The Near-Infrared Camera and Multi-Object Spectrometer Ultra Deep Field: Observations, Data Reduction, and Galaxy Photometry. *AJ*, 130(1):1–12.

Weinberg, S. (1972). *Gravitation and cosmology*. Wiley.

Windhorst, R. A., Cohen, S. H., Hathi, N. P., McCarthy, P. J., Ryan, Russell E., J., Yan, H., Baldry, I. K., Driver, S. P., Frogel, J. A., Hill, D. T., Kelvin, L. S., Koekemoer, A. M., Mechtley, M., O'Connell, R. W., Robotham, A. S. G., Rutkowski, M. J., Seibert, M., Straughn, A. N., Tuffs, R. J., Balick, B., Bond, H. E., Bushouse, H., Calzetti, D., Crockett, M., Disney, M. J., Dopita, M. A., Hall, D. N. B., Holtzman, J. A., Kaviraj, S., Kimble, R. A., MacKenty, J. W., Mutchler, M., Paresce, F., Saha, A., Silk, J. I., Trauger, J. T., Walker, A. R., Whitmore, B. C., and Young, E. T. (2011). The Hubble Space Telescope Wide Field Camera 3 Early Release Science Data: Panchromatic Faint Object Counts for 0.2-2 μm Wavelength. *APJS*, 193(2):27.

Windhorst, R. A., Timmes, F. X., Wyithe, J. S. B., Alpaslan, M., Andrews, S. K., Coe, D., Diego, J. M., Dijkstra, M., Driver, S. P., Kelly, P. L., and Kim, D. (2018). On the Observability of Individual Population III Stars and Their Stellar-mass Black Hole Accretion Disks through Cluster Caustic Transits. *APJS*, 234(2):41.

Wright, E. L. (2006). A Cosmology Calculator for the World Wide Web. *PASP*, 118(850):1711–1715.

Yan, H. and Windhorst, R. A. (2004). Candidates of z $\sim$= 5.5-7 Galaxies in the Hubble Space Telescope Ultra Deep Field. *APJL*, 612(2):L93–L96.

13 Appendix A. Derivation of Equations

Here we include our derivations of the math used in *AHaH* at a level appropriate for an intermediate undergraduate cosmology class[10].

13.1 Comoving Radial Distance

To begin, we need the comoving radial distance, D_R, from the Earth to an object at redshift z, derived from the Robertson-Walker metric, as discussed previously, *e.g.*, Longair (Ch. 7 1998), Eqs. 5.33 and 6.13 of Ryden (2017), and Eq. 6 of Wright (2006). We express this as the integral:

$$D_R(z) = \int_t^{t_0} \frac{c \cdot dt}{a} = \int_{\frac{1}{1+z}}^{1} \frac{c \cdot da}{a\dot{a}} = \frac{c}{H_0}\int_0^z \frac{dz}{(1+z)\dot{a}}, \qquad (5)$$

where the scale factor $a = 1/(1+z)$. The derivative of a with respect to time, $\dot{a}$, is given by the expression:

$$\dot{a} = (\Omega_M/a + \Omega_R/a^2 + \Omega_\Lambda \cdot a^2 + \Omega_K)^{1/2}, \qquad (6)$$

where Ω_M, Ω_R, Ω_Λ, and Ω_K are energy density parameters, corresponding to the fractions of the Universe's total average energy density that are attributable to matter (M), radiation (R), dark energy (Λ), and the curvature of the spatial geometry (K), respectively. Note that it is assumed these are the only relevant contributions to the total energy density Ω_{Tot}. That is, we assume that $\Omega_{Tot} = \Omega_M + \Omega_\Lambda + \Omega_R + \Omega_K$. A spatially flat Universe would have $\Omega_{Tot} \equiv 1$ with $\Omega_K = 0$. The default Planck Collaboration et al. (2018) parameters used are: H_o = 68 km/sec/Mpc, $\Omega_M = 0.32$, $\Omega_\Lambda = 0.68$, $\Omega_R = 9\times10^{-5}$ with $\Omega_K = 0$.

We evaluate this integral in steps of 0.05 in z from $z = 0$ to $z = 20$ to create a look-up table, interpolating linearly to find the value for any arbitrary redshift in between these discrete steps. This is because we must make the calculation frequently and for many objects, so computing the integral manually every time would be computationally prohibitive. The resulting error in this method is generally small enough that it translates to less than one pixel's difference even on high-resolution displays, so this error can safely be ignored for the purposes of the application. We evaluate the integral using the simple midpoint method, which may not be the optimal solution, but was simple to implement and adequately efficient on any home computer. As with the linear interpolation, higher accuracy numerical integration would result in less than one pixel's difference when displayed.

13.2 Angular Size Distance

To develop the angular size distance, D_A, we first need to develop a generalized form of the comoving radial distance D_R, to express the distance measure to an object at redshift z_j as measured by an observer at redshift z_i. This distance is given by the formula:

$$D_R(z_i, z_j) = \begin{cases} \mathfrak{R}_i \sin(r_j/\mathfrak{R}_i) & \text{if } \Omega_K < 0 \\ r_i & \text{if } \Omega_K = 0 \\ \mathfrak{R}_i \sinh(r_j/\mathfrak{R}_i) & \text{if } \Omega_K > 0 \end{cases}, \qquad (7)$$

where $\mathfrak{R}'$ is the radius of curvature of the spatial geometry at redshift z_i, and r_i is the value of the comoving coordinate distance at the same redshift (Longair, 1998; Wright, 2006). These

10 See, *e.g.*, `http://windhorst322.asu.edu`

correspond to the cases where the spatial geometry of the Universe is spherically curved, flat, and hyperbolically curved, respectively. Recalling that both r'_{ij} and $\mathfrak{R}'$ scale as $1/(1+z_i)$, and that $\mathfrak{R} = (c/H_0)/\sqrt{|\Omega_K|}$, we next define an intermediate quantity U, representing the argument of sin and sinh in Equation (7) above:

$$U = r_i/\mathfrak{R}_i = r_0/\mathfrak{R}_0 = (H_0/c)\sqrt{|\Omega_K|}r_0 \tag{8}$$

We note that since U now depends only upon the cosmology selected by the user and the object's redshift, we may calculate U once per object and re-use it, thus saving CPU time. Using this quantity, we may now rewrite D_R as:

$$D_R(z_i, z_j) = \frac{\delta(U)}{1+z_i} r_0 \tag{9}$$

Here, $\delta(U)$ is simply some function of U. By substituting U into Equation (7) above, we get the following expression for $\delta(U)$:

$$\delta(U) = \begin{cases} \frac{\sin(U)}{U} & \text{if } \Omega_K < 0 \\ 1 & \text{if } \Omega_K = 0 \\ \frac{\sinh(U)}{U} & \text{if } \Omega_K > 0 \end{cases} \tag{10}$$

Note that $\delta(U)$ expressly depends upon r_0. The case where $\Omega_K = 0$ comes from the limit of both $\sin(U)/U$ and $\sinh(U)/U$ as $\Omega_K \to 0$ – one may observe that in this case Equation (9) simplifies to $r_0/(1+z_i)$, which is precisely r_i as in Equation (7), since as observers we start our *AHaH* journey at $z_i = 0$. When we expand with *AHaH* into the HUDF galaxy database as it is sorted in redshift, the observer's redshift can take any value $0 \lesssim z_i \lesssim 20$, although the current set of HUDF postage stamp images does not have any galaxies with $z \gtrsim 6$. We intend to expand this with a future version of *AHaH* based on the full data set summarized in Windhorst et al. (2011).

Thus, using our Equation (9) and the equation relating the angular size distance and the distance measure as developed by Longair (1998, Eq. 7.50), the angular size distance from redshift z_i to z_j is given by:

$$D_A(z_i, z_j) = D_R(z_i, z_j)\frac{1+z_i}{1+z_j} = \frac{\delta(U)}{1+z_j} r_0 \tag{11}$$

13.3 Comoving Coordinate System

Now that we have developed formulae for D_R and D_A, we can consider the best way to create a coordinate system for the Java application. The data we start with are the redshift z_j of object j (with which we can calculate D_R) and four angles: the object's angular size (from the height and width of its image) and the angular separation between the object and the x and y axes, which we define as lines going through the center of the original image. These angles are calculated by taking the corresponding size in pixels and multiplying by the scale in arcsec/pixel of the original HST image[11].

We would like to use this information to create a coordinate system with the original telescope position at the origin. In a Euclidean space this would present no problem, but we have already remarked that the *observed* angles are *not* the same in an expanding Universe as they would be in a Euclidean space. Further, it would be desirable for the Euclidean coordinate distance to correspond to the comoving radial distance, as this would make calculations significantly simpler. We can accomplish this, but when we create coordinates for each object as such, we need to "correct" the observed angles. That is, we want a "Euclidean angular size" associated with a certain observed angular size. We will call this θ_E. An object's angular size is related to its physical transverse diameter, d, by the following equation which derives from Equation (11):

$$d = \theta D_A = \theta\frac{\delta(U)}{1+z_j} r_0 = \theta_E \frac{1}{1+z_i} r_0 \tag{12}$$

Note that in the Euclidean case we must contract r_0 by a factor of $1/(1+z_i)$ to get the comoving distance from z_i to z_j as measured by the observer at z_i (r_i in Equation (7) above). This is because the proper spatial separation in the current epoch has been stretched by the Universe's expansion, so as seen by an observer at redshift z_i it must be scaled appropriately. Hence, the equivalent Euclidean distance between any two points is $D_E = r_0/(1+z_i)$, in which case we get the Euclidean small-angle approximation back: $d = \theta_E D_E$.

Thus canceling r_0, we get the following expression for θ_E from Equation (12):

$$\theta_E = \theta\delta(U)\frac{1+z_i}{1+z_j} \tag{13}$$

In our initial data z_i is simply zero, so that we create coordinates (X, Y, Z) for an object at redshift z like:

$$X = \sin\left(\frac{\delta(U)\theta_X}{1+z}\right)\cos\left(\frac{\delta(U)\theta_Y}{1+z}\right) D_R(0, z), \tag{14}$$

and similarly for Y and Z. We have thus developed a coordinate system of X, Y, and Z in comoving Mpc with the original telescope position at the origin.

13.4 Simulating Observations From Vantage Points Other Than $z = 0$

Now, when we "move" the Hubble camera virtually to higher redshifts, we do so by moving to some new (X_C, Y_C, Z_C) value in the coordinate space. (Note that *AHaH* does not only "virtually violate the laws of physics" by moving the observer into the HUDF images at $\sim 500 \times 10^{12}$ times the speed of light, but it also has no problem "violating the arrow of time" by allowing the user to move back and forth in redshift through the sorted HUDF image data cube). By construction, the distance measure here is just the Euclidean coordinate distance:

$$D_E = ((X - X_C)^2 + (Y - Y_C)^2 + (Z - Z_C)^2)^{1/2} \tag{15}$$

Now to determine where to display an object after we have "moved" the camera, we use the distance calculated with Equation (15) and the Euclidean angular size. Using the redshift of the object of interest, z_O, and the camera's user-defined redshift, z_C, we rearrange Equation (13) to get:

$$\theta = \theta_E \frac{1+z_O}{\delta_{ij}(1+z_C)} \tag{16}$$

Here δ_{ij} follows from Equation (7) and Equation (10) for an object at redshift z_j as observed from redshift z_i. In this case, θ_E is a quantity that we must calculate from our coordinates in the

11 The HUDF mosaics used have been drizzled to a scale of 0.″03 per pixel

usual Euclidean way.

For an object's angular size it is even simpler than for its (X, Y, Z) position, since we do not have to manually calculate θ_E. We know that in the Euclidean case:

$$d = \theta_0 D_R = \theta_E D_E, \tag{17}$$

where θ_0 is the Euclidean angular size from redshift zero, and D_E is the coordinate distance from the camera to the object from Equation (15). We then solve for θ_E in Equation (17), $\theta_E = \theta_0 D_R / D_E$, and substitute back into Equation (16) to obtain an expression for the desired angular size θ for any object at redshift z_o as observed from z_c:

$$\theta = \theta_0 \left(\frac{D_R}{D_E} \right) \frac{1 + z_o}{\delta_{ij}(1 + z_c)} \tag{18}$$

14 Appendix B. Specialized Data Preparation

To develop the images used in *AHaH* we first created a custom-balanced RGB version of the HUDF image[12]. While the image provided in the original press releases[12] would have been adequate, it has the undesirable characteristic that very bright areas, such as bulges in large spiral galaxies, appear burned-out and lack fine detail. The HUDF data used was taken to roughly equal depths in the B-, V-, i'-, and z'-band filters, which have central wavelengths of ~4320 (Blue), 5920 (Visual), 7690 (Red), and 9030 Å (near-Infrared), respectively (Beckwith et al., 2006), so we created a three-channel color image by first combining the B- and V-bands, applying weights based on the sky signal-to-noise ratio [13]. We then used the algorithm developed by Lupton et al. (2004) to create the combined RGB image, with the combined $B+V$-bands as the blue channel, the i'-band as the green channel, and the z'-band as the red channel[14]. Besides showing more detail in bright areas, this method has the added benefit that an object with a specified astronomical color has a unique color in the composite RGB image. A comparison of the original STScI color images and our prepared images is shown in Figure 1. The full HUDF image using this color preparation technique is also available as an interactive map online[15]. One can see examples of galaxies processed in this way in Figure 2.

The galaxies represented in the *AHaH* application were i'-band (8000 Å) selected using the `SourceExtractor` algorithm (Bertin and Arnouts, 1996; Bertin et al., 2002) with a detection threshold of 1.8× the rms noise level (1.8σ) above the local sky. The i'-band (at $z \approx 6$) dropouts of Yan and Windhorst (2004) were added by hand. We then created color JPEG "stamp" images for each individual object, using the `SourceExtractor`-generated segmentation map to mask as black any pixels outside the detected source. These "stamps" were then converted pixel-for-pixel to PNG images, which employ a lossless compression algorithm – no image quality was thus lost. We then developed a transparency map based on each pixel's brightness, which was saved into the PNG alpha channel[16]. The resulting images can thus be displayed as *semi-transparent*, allowing objects in the distance to show through the dim regions of objects in the foreground, as is also possible in the real Universe.

Photometric redshifts for the galaxies were measured with the `HyperZ` package (Bolzonella et al., 2000), using a combination of the original HST-ACS four-band ($BVi'z'$) data from the HUDF, along with Y-, J- and H-band near-IR data from HST NICMOS (Near Infrared Camera and Multi-Object Spectrometer) (Thompson et al., 2005). We have supplemented the photometric redshifts with spectro-photometric redshifts measured by Ryan et al. (2007), which incorporate the aforementioned $BVi'z'JH$ data, as well as grism spectra from GRAPES (Pirzkal et al., 2004, 2017), the U-band observations from CTIO (Cerro Tololo Inter-American Observatory) Mosaic II (Dahlen et al., 2007), and K_S-band data from VLT-ISAAC (Very Large Telescope – Infrared Spectrometer And Array Camera, e.g. Retzlaff et al. (2010)). For a summary of all these data and the quality of the spectro-photometric redshifts, see Ryan et al. (2007). When available, we have chosen to use the more reliable spectro-photometric redshifts.

12 See `http://imgsrc.hubblesite.org/hu/db/2004/07/images/a/formats/full_jpg.jpg` for the full-resolution 60 Mb HUDF ACS image in the BViz filters, and `https://www.asu.edu/clas/hst/www/aas2014/HUDF14-pan-UVrendered.jpg` for the deepest 13-filter panchromatic 841-orbit Hubble image.

13 We applied a weight of 0.765 in V and 0.235 in B to balance the image depths. For details, see Windhorst et al. (2011)

14 The channels were first scaled proportional to the data zero points – Red: 716.474, Green: 345.462, Blue: 254.449, see `https://hst-docs.stsci.edu/acsihb`.

15 `http://ahah.asu.edu/clickonHUDF/index.html`

16 The alpha channel of a PNG contains the "transparency" of each pixel, which is more straightforwardly the opacity of the pixel. For more, see the PNG specification page from w3: `https://www.w3.org/TR/PNG-DataRep.html`.

ASTRONOMY EDUCATION JOURNAL

Astronomy Education Journal, 2021, Vol. 01, No. 1

doi: 10.32374/AEJ.2021.1.1.004

RESOURCES & ACTIVITIES

A historical method approach to teaching Kepler's 2nd law

Wladimir Lyra[1,*]

[1]New Mexico State University, USA

[*]wlyra@nmsu.edu

Abstract

Kepler's 2nd law, the law of the areas, is usually taught in passing, between the 1st and the 3rd laws, to be explained "later on" as a consequence of angular momentum conservation. The 1st and 3rd laws receive the bulk of attention; the 1st law because of the paradigm-shift significance in overhauling the previous circular models with epicycles of both Ptolemy and Copernicus, the 3rd because of its convenience to the standard curriculum in having a simple mathematical statement that allows for quantitative homework assignments and exams. In this work I advance a method for teaching the 2nd law that combines the paradigm-shift significance of the 1st law and the mathematical proclivity of the 3rd law. The approach is rooted in variational learning and the historical method, indeed, placed in its historical context, Kepler's 2nd is as revolutionary as the 1st: as the 1st law does away with the epicycle, the 2nd law does away with the equant. This way of teaching the 2nd law also formulates the "time=area" statement quantitatively, in the way of Kepler's equation, $M = E - e\sin E$, (relating mean anomaly M, eccentric anomaly E, and eccentricity e), where the left-hand side is time and the right-hand side is area. In doing so, it naturally paves the way to finishing the module with an active learning computational exercise, for instance, to calculate the timing and location of Mars' next opposition. This method is partially based on Kepler's original thought, and should thus best be applied to research-oriented students, such as junior and senior physics/astronomy undergraduates, or graduate students.

Keywords: Culture-Based Astronomy Education Research; Astronomy Laboratory Activities; Kepler's Laws; Historical Method; Variational Theory of Learning.

1 Introduction

Kepler's 2nd law is arguably the most challenging of Kepler's laws to teach. Yu et al. (2010) found that, in a sample of 112 undergraduate student interviews to gauge prior knowledge for an introductory astronomy course, the majority (54%) declined to even guess an answer when inquired about it. This contrasts to the 1st and 3rd laws, where a majority of the same students gave incomplete but correct statements about them ("orbits are not circles", and "planets orbiting closer to the Sun move faster; those orbiting farther move slower", respectively). While the sample of Yu et al. (2010) was of nonmajor freshmen, the 2nd law remains underrated in upper division major undergraduate courses, where students' understanding of it still lingers on the qualitative, and divorced from its historical significance. In addition, Aktan and Dinçer (2014) find alternative conceptions about the 2nd law even among pre-service science teachers. This evidences shortcomings about the way the law is traditionally explained. The 2nd law is frequently taught (e.g.: Carroll and Ostlie, 2007; Ryden and Peterson, 2020) as a variation of the following statement:

> *A line drawn from the Sun to a planet sweeps out equal areas in equal time intervals.*

This sentence is usually followed by diagrams showing short wide areas near the Sun and long slender areas far away. "They're equal!", says the instructor, almost like a curiosity. "Why is this important?", asks the inquisitive student. Planets go fast near the Sun and slow far from it, is the usual answer. It is how it is described in popular astronomy books (e.g.: de Cayeux and Brunier, 1983), in high school physics (e.g.: Kuhn, 1962; Guimaraes and Fonte Boa, 2006), and in numerous educational websites. In college, one walks the extra mile of proving from Newton's laws that Kepler's 2nd simply reflects angular momentum conservation (e.g.: Halliday and Resnick, 2020; Marion and Thornton, 2020). This is usually done by climbing down the pedestal of differential and integral calculus to the pedestrian world of Euclidian geometry and defining the strange concept of "areal velocity", which is then proved to be constant.

The historical inversion (Newton before Kepler) is rooted on a discipline approach (Rice, 1972). By grounding Kepler's 2nd law on angular momentum conservation, it draws on principled conceptual knowledge (Leinhardt, 1988), facilitating learning by structuring it around a major concept of the discipline of physics. On the other hand, teaching the material in this way has the unfortunate drawback of reducing Kepler's 2nd law to a post-factum instead of presenting it as the product of original logic, painstaking problem-solving, and what was then cutting-edge research. These, in turn, are precisely the skills that should be developed in high-ability students (Dixon et al., 2004). More importantly, this presentation fails to obviate that the 2nd law is about *quantitatively* finding the planet in the orbit. In this paper I develop a model for teaching Kepler's second law using the historical method (Matthews, 1989; Coelho, 2009; Galili, 2010). The proposed teaching sequence partially recreates in the classroom the historical perspective in which Kepler discovered the law in *Astronomia Nova* (Kepler, 1609; Aiton, 1969; Boccaletti, 2001).

2 The Historical Method

The historical method (also called genetic approach) has a rich record in physics. Mach (1895, 1911) and Duhem (1906) argued that retracing the original line of thought in discovering the laws of nature led to a deeper understanding of the subject by the novice students. The sentiment is echoed by Schwab (1962), who defined teaching *science as inquiry* in its essence as "to show some of the conclusions of science in the framework of the way they arise and are tested". Modern pedagogy frames this postulate under the idea of cognitive recapitulation (Piaget, 1970; Posner et al., 1982): ontogeny recapitulates phylogeny, i.e. there is a parallel between how an individual accrues knowledge (ontogeny) and how the knowledge in the discipline itself evolved (phylogeny).

Indeed, the way by which conceptual change is brought about in learners (Posner et al., 1982) shares many similarities with the structure of paradigm shifts (Kuhn, 1962). According to Posner et al. (1982), learners have no urge to change their conception until that conception fails at problem solving. This is how scientific theories are constructed, thriving until anomalous data is introduced, that cannot be accounted for by the prevailing model. Also, the first response of learners to anomalous data is not dismissal of the old conception, but questioning of the data. This is mirrored by the healthy skepticism of the scientific community, requiring extraordinary claims to be backed by extraordinary evidence. Finally, once the quality of the new data is established, learners arrive at a conceptual change when presented with a new theory that accounts for both the old and the new data. This is epitomized by the correspondence principle (Bohr, 1920), which requires a new theory to explain all the phenomena for which a preceding theory was valid. Visibly, the historical method retraces the development of the field, prompting the student to understand under what circumstances the previous theory was judged convincing, and why a new theory is necessary.

Equally important, by recreating the atmosphere of discovery, the historical method inherently brings into the classroom the culture of the field (Conant, 1964; Holton, 1978), informing not only knowledge but also its structure: the deductive logic and reasoning by which knowledge was constructed, older and now obsolete principles and concepts, how they were replaced, the thinking of generations of astronomers – Galili (2010) calls this the *periphery* of the discipline, as opposed to the discipline's nucleus (axioms and laws) and body (applications).

Finally, research in human cognition shows that learning is facilitated when presented with a contrast between alternatives (Piaget, 1977; Zazkis and Chernoff, 2006; Mansouri et al., 2009; Waxer and Morton, 2012), which is explained within the framework of Variation Theory of Learning (Marton and Booth, 1997; Pang and Marton, 2003; Marton et al., 2004; Marton and Pang, 2008; Orgill, 2012; Bussey et al., 2013; Cheng, 2016). According to variation theory, for learning to occur, some critical aspects of the object of learning must vary while other aspects remain constant. This is exactly how scientists determine cause and effect, mapping the parameter space by exploring one variable at a time while holding all the other variables constant. As a consequence, another reason the historical method is effective as a teaching tool is because, in recreating the narrative that led to the production of knowledge, it recreates the conceptual conflict that necessitated that knowledge (Monk and Osborne, 1997), while also using the tools of the scientific method adapted to the classroom in the form of variational learning.

This paper is structured as follows. In the next section the context of the study is presented, followed by the teaching design (the proposed teaching sequence), including exposition and active learning exercises in the presentation for replicating the method. I conclude with an assessment of student learning and discussion.

3 Context of the Study

This method was originally developed as part of a one semester course on dynamical astronomy for physics students in their junior year, at a primarily undergraduate university in California, in 2018. The class had 24 students. I taught it again twice for first-year astronomy PhD students (ASTR 503, "Fundamental Astronomy") at New Mexico State University, in 2019, and 2020. I teach this in two classes of 75 minutes each, as part of a module on Kepler's laws. The first two times were taught in person, the third time online in "flipped classroom" format (King, 1993; the videos are available at `https://www.youtube.com/playlist?list=PLatuGW739E0lVsAwwqTHKU0tluD9h4c3I`). To com-plement the pre-class videos, typeset notes are also provided (available at `http://astronomy.nmsu.edu/wlyra/FundamentalAstronomy/Module3_KeplerLaws_Notes.pdf`). The 2019 class had 10 students, which provided a convenience sample (Saumure and Given, 2008) to poll about the 2nd law.

The fundamental question that guides the module is a question that intrigued humanity for millenia: *how to predict the position of the planets?* The goal of the module is to understand how Kepler's laws connect to the emergence of modern astronomy, to understand planetary motion, the interplay between theory and observations, and the fundamental importance of observational accuracy. The module on Kepler's laws is done after a module on Spherical Astronomy, so the students are familiar with coordinate systems on the celestial sphere, and how to transform between equatorial and ecliptic coordinates. I also introduce the concept of elongation, the angle between the planet and the Sun. This serves the purpose of introducing Ptolemy's model, which is key to understanding

Table 1. Initial conditions for the N-body code for computing the position and velocities of "real" Earth and Mars.

Planet	Semimajor axis (a)	Eccentricity (e)	Inclination (i)	Position (xyz)	Velocity (v_x, v_y, v_z)
Earth	1.0000	0.0000	0	(1,0,0)	(0,1,0)
Mars	1.5237	0.0934	0	(1.66601358,0,0)	(0, 0.73768098,0)

the revolutionary character of Kepler's 2nd law. For the junior class I started with Owen Gingerich's celebrated Mars lab (Gingerich, 1983) to find Kepler's 1st law in an active learning way (Bonwell and Eison, 1991). For the PhD course, Gingerich's lab was done as a computational exercise. The lab is followed by the geometrical proof that the orbit is an ellipse, again using the historical method, with pre-Newtonian reasoning (Appendix A). In the online version I could not provide drafting tools to each student, so I created a video of the method (available at `https://www.youtube.com/watch?v=Ss-nmWFY5Wo&list=PLatuGW739E01VsAwwqTHKU0tluD9h4c3I&index=2`), and we did active learning in class showing how we would not be able to discriminate between an ellipse and an off-centered circle with the accuracy of common classroom drafting instruments. The 1st law lab and instruction set the stage for Kepler's 2nd law.

4 Teaching Methods

The historical method was adopted in response to the unsettledness of teaching Kepler's 2nd Law as a post-factum, and with seemingly less importance than the 1st and the 3rd. The 1st law, *the planets orbit the Sun in elliptical orbits with the Sun at one of the two foci*, has a clear and powerful paradigm-shifting formulation. Its statement is a direct and unequivocal breaking with the previous cosmological models, of both Ptolemy and Copernicus, that insisted on circular orbits. The 3rd law, *the cube of the semimajor axes is proportional to the square of the periods*, is formulated as an elementary mathematical statement, and thus conveniently translated into quantitative homework assignments and exams, even at the high school level. In contrast with the 1st and the 3rd, the second law, *the radius vector connecting the planet to the Sun sweeps equal areas in equal times*, sounds disturbingly turbid to the modern student, its geometric statement a remnant of a pre-calculus era. Transplanting ourselves to Kepler's time by putting aside Newtonian physics and knowledge of conservation of angular momentum, one should ask: *why did Kepler care about area?* In the junior class in 2018 I asked the students this question. No hands were raised. Not wanting to repeat the usual way of teaching the 2nd law, and given how other educators also struggle with how to present it (Setyadin et al., 2020), I decided to teach it partially following the historical method to retrace Kepler's original line of thought. In its historical context, Kepler 2nd law is similar in formulation to the 1st law in the sense that it is contrasted to the previous model. *The orbit is an ellipse contrasts to the orbit is a circle. Likewise, equal areas at equal times* contrasts to equal angles at equal times. The 2nd law is formulated as a conceptual conflict. For 1500 years, up to Kepler, astronomy insisted not only on circles but also on uniform motion. In Ptolemy's model, to account for the perceived non-uniform motion of a planet, he introduced the equant, which is a point on the line of apsides about which the center of the epicycle does uniform motion. The practicality is that time is given by angle, so the motion of the planet, though non-uniform from Earth's reference frame, is easily parametrized in time. The doctrine of uniform motion was so prevalent that it was an *a priori* in Copernicus new heliocentric model: to do away with non-uniform motion and cast all planetary movement as uniform circular motion in deferents and epicycles about their centers. Astronomy up to Kepler had it ingrained that there was a reference frame about which a planet sweeps equal angles at equal times. That is the prior model that the 2nd law contests. Imagine that you never heard about circular orbits before. It would become difficult to understand the paradigm-shifting impact of the 1st law. This lack of awareness is precisely the situation a modern student encounters the 2nd law.

The remainder of this section will be presented for a flipped classroom (King, 1993) format. In this style, content delivery is removed from the in-person classroom time with the students, which is used instead for discussion and active learning. The method here developed to teach Kepler's 2nd law shares similarities with the presentation of the same subject by Holton and Brush (1952). Our method, however, is more mathematically grounded and focuses on the correspondence principle and conceptual conflict between the equant model and Kepler's 2nd law. A drawback of the method is the need to teach the equant model, which many students (as well as instructors) have little familiarity with. However, we minimized the time needed to introduce the model, while also keeping it pedagogical.

4.1 Presentation of the teaching sequence

Finding the shape of the orbit is not solving the whole problem of planetary motion. A practical question remains: how to find the planet in the orbit? To better understand Kepler's 2nd law, let us look at what existed before him. Ancient wisdom insisted in uniform circular motion, because it was their way to understand periodicity. Regularity was found in circles, according to Copernicus (1543), *"the only figure that can bring back the past"*. Although it was apparent that the Sun (or the Earth for that matter) was not the center of the orbit, the shape of the ellipse was out of the reach of their observational accuracy. Non-uniform motion along the orbit was also evident, and a solution was found by Ptolemy, namely, the equant model. Most modern astronomers are not familiar with this model, which is nowadays only of historical importance, so a brief pedagogical exposition is warranted. Let us get to it step by step. Along the way, the students will do active learning exercises to understand Kepler's 2nd law. Like Kepler, we will use Mars in this study, because it is the superior planet of highest eccentricity and, added bonus, also the closest one.

For the exercises presented in this section, the students will use a N-body code of their choice – e.g., Rebound (Rein and Liu, 2012) or Mercury (Chambers, 2012); the plots here shown were calculated with the Pencil Code (Brandenburg et al., 2021) – and calculate the orbital evolution of the Earth and Mars. Earth and Mars are test particles (zero mass), so the Sun's position coincides with the barycenter of the system. The results here shown have Earth and Mars initialized at the position and velocities given in Table 1, which are the position and velocity at aphelion. The orbit of the Earth is approximated to zero eccentricity, and the orbits of both planets to zero inclination. In these units an Earth year is $T = 2\pi$. Run the simulation until time t=100, which is approximately 16 Earth years. The students will use this "real" Mars to compare to the different models.

The critical aspect that the instructor should make

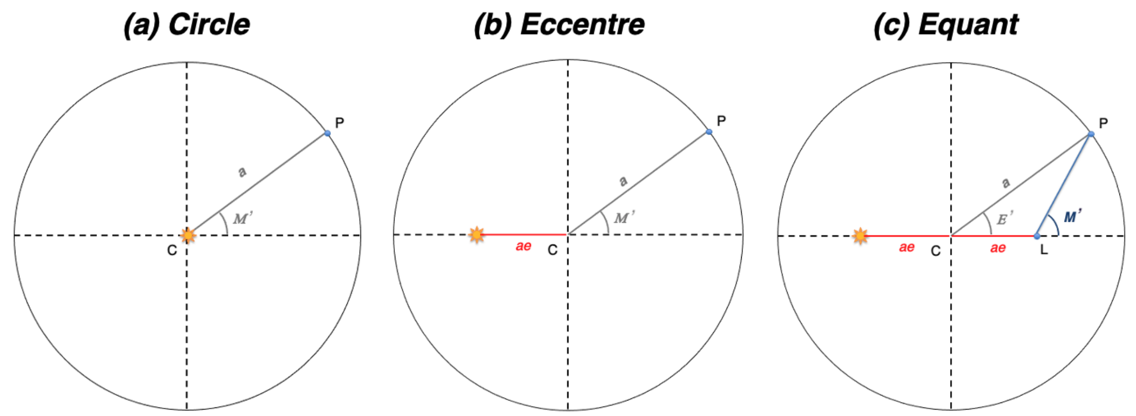

Figure 1. Elements of the different circular orbit models: (a) Sun-centered uniform circular motion about the Sun; (b) Off-centered uniform circular motion about the center; (c) Off-centered uniform motion about the equant (point L). *M* is the mean anomaly and *E* the eccentric anomaly. On (a) and (b) the mean anomaly and the eccentric anomaly are identical.

the students aware of is that the accuracy of ancient observations was 1°, so any model that fits the positions of the planets to one degree accuracy will be deemed acceptable. Otherwise, the model is rejected.

4.2 Starting simple: a sun-centered, circular model

Let us assume that planets go in circular orbits, centered at the Sun, in uniform motion (Fig. 1a). At any instant of time, the position of Mars is given by:

$$x(t) = a \cos M'(t) \quad (1)$$
$$y(t) = a \sin M'(t) \quad (2)$$

Here M' is measured counterclockwise from aphelion; it relates to the mean anomaly M (measured counterclockwise from perihelion) by $M' = M - 180$. The mean anomaly is $M(t) = nt$, with t meaning time, and $n = \frac{2\pi}{T}$ is the mean motion, where T is the period of Mars. Here, M, the mean anomaly, is equal to both the eccentric and true anomalies. Another critical aspect to raise awareness here is that everything but time in the definition of mean anomaly is constant. Mean anomaly equals time. *Mean anomaly is time*.

The students will compare this sun-centred circular model with the "real" Mars they calculated from the N-body model. The students should plot the ecliptic longitude of Mars in the sky, which is:

$$\lambda = \arctan\left(\frac{\Delta y}{\Delta x}\right) \quad (3)$$

Where $\Delta y = y(t)_{Mars} - y(t)_{Earth}, and \Delta x = x(t)_{Mars} - x(t)_{Earth}$; i.e. the relative position of Mars and the Earth. The students should plot the four subplots of Figure 2. The upper left plot is λ vs time for the two models ("real" Mars in red and the circular model of Eqs 1 and 2 in black). The upper right plot is the deviation in λ between the model and real Mars. The lower left plot is the bird-eye heliocentric view of the orbits, i.e., $y(t)$ vs $x(t)$. The lower right plot is the bird-eye geocentric view, i.e., Δy vs Δx.

From this exercise, the students should realize that the model does not reproduce either the shape of the orbit or the longitudes. The predicted positions of retrogradations, specifically, are off by as much as 30 degrees from the actual positions of Mars. The students here discern that the circle model cannot be correct, and conclude that the model has to be discarded. The instructor can now introduce variation, fitting other models, and performing the same analysis to assess the adequacy and accuracy of each model in reproducing the observations.

4.3 The Eccentre

The circular model failing, the instructor now introduces a new model, the eccentre (Fig. 1b). The critical aspect is centredness: this model merely shifts the position of the center of the orbit away from the Sun by an amount ae, keeping the uniform motion. At any given instant in time, the position of Mars, seen from the Sun, is now given by:

$$x(t) = a \cos E'(t) + ae \quad (4)$$
$$y(t) = a \sin E'(t) \quad (5)$$

where ae is the amount we shift the center away from the Sun. Here E', like M', is measured counterclockwise from aphelion; it relates to the eccentric anomaly E (measured counterclockwise from perihelion) by $E' = E - 180$.

The students should compare this model with the "real" Mars. They will plot the same graphs as Figure 2, but now with this model. The results are seen in Figure 3, where the model is shown in cyan. The students realize that the model reproduces the orbit, but it does not reproduce the velocity of Mars. It predicts oppositions and retrogradations still off by 15 degrees. Again, this model cannot be right, and the students discard the model.

4.4 Non-uniform motion

Another variation will be introduced, keeping the eccentre, since it reproduces the orbit, but relaxing the idea of uniform motion in order to reproduce the orbital velocity. The instructor now finally presents the equant model (Fig. 1c). To fit the velocity of the planets, Ptolemy added a third device, the equant point, defined as a point on the line of apsides about which the angular velocity of a body on its orbit is constant. This point is point L in Figure 1c. About L, the planet, located at P, goes around in uniform motion, being described by the angle $M = nt$. The eccentric anomaly E seen from the center of

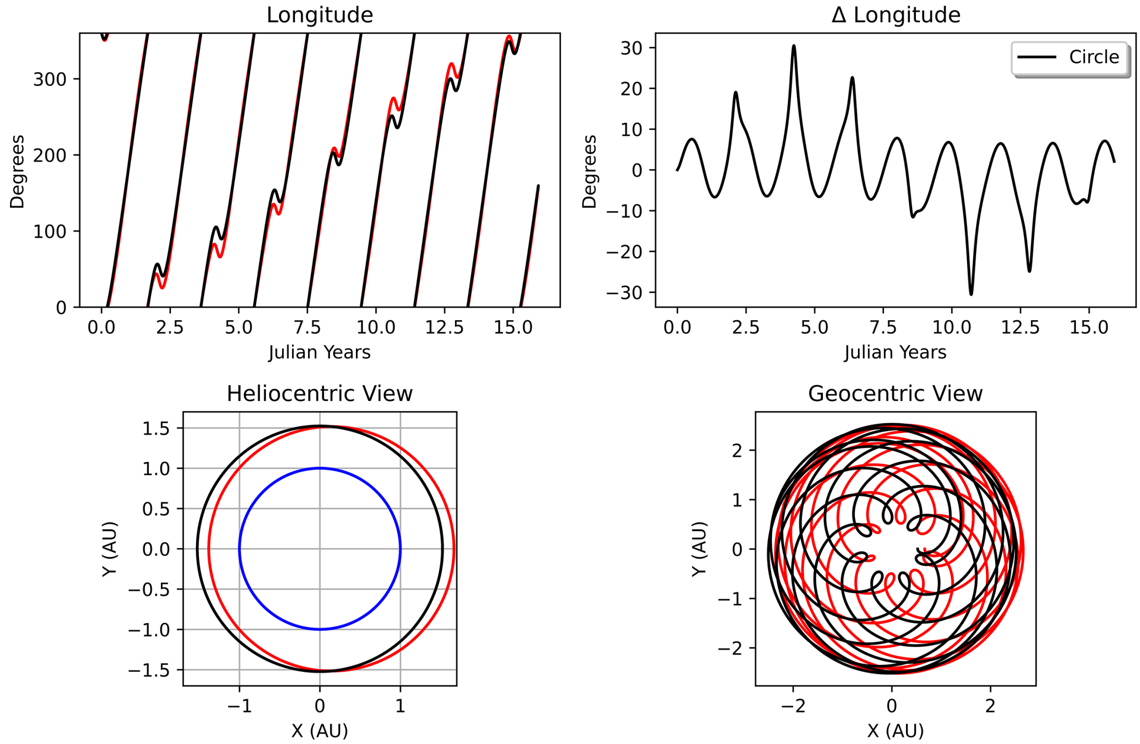

Figure 2. Mars orbit versus a circular Sun-centered orbit model. Upper left plot: Ecliptic longitude vs Time, Mars (red) vs circular orbit (black). The longitudes generally match, except at retrogradations. Upper right plot: Longitude residual. The error amounts to as much as 30 degrees. The circular Sun-centered model is not acceptable. Lower left plot: Heliocentric view of the orbit. Red is mars, blue is Earth, magenta the model. Lower right plot: Geocentric view of the orbit.

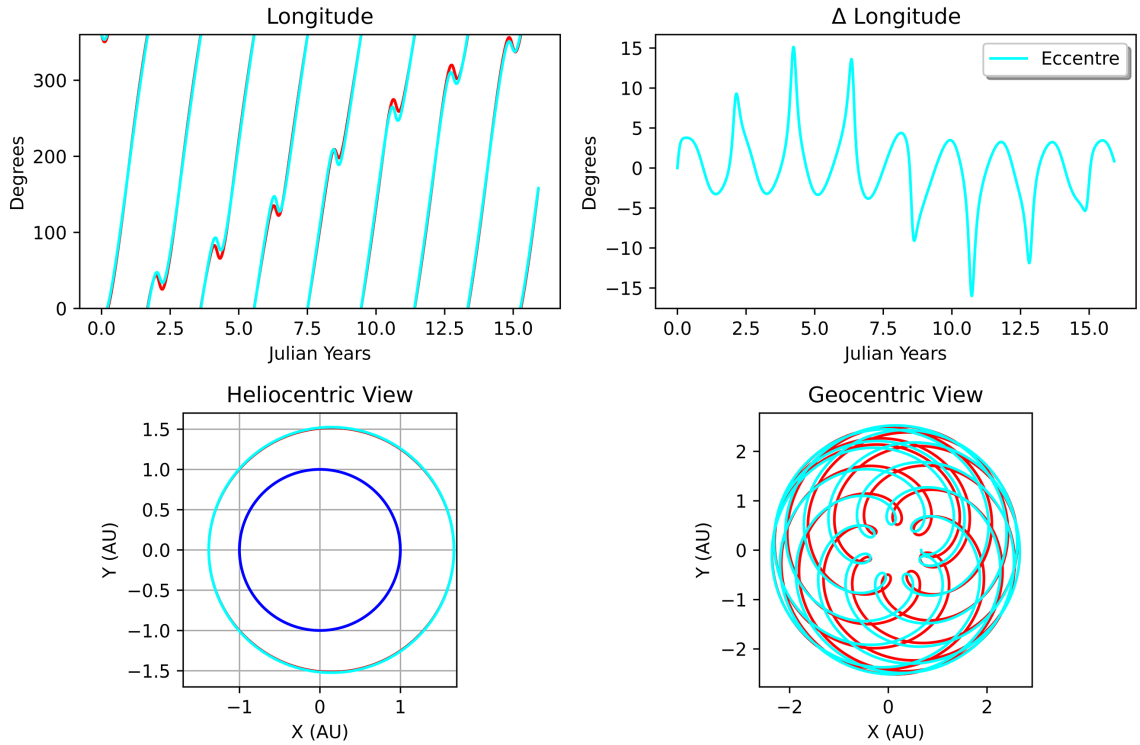

Figure 3. Mars orbit versus off-centered circular orbit model (the eccentre), keeping uniform motion. Upper left plot: Ecliptic longitude vs Time, Mars (red) vs circular off-centered orbit (cyan). The longitudes generally match, except at retrogradations. Upper right plot: Longitude residual. The error is better than the Sun-centered model, but still amounts to as much as 15 degrees. The off-centered model with uniform circular motion is not acceptable. Lower left plot: Heliocentric view of the orbit. Red is mars, blue is Earth, green the model. Lower right plot: Geocentric view of the orbit.

Table 2. Initial conditions for the N-body code for computing the evolution seen in Fig 5. All planets are massless. The other orbital elements (inclination, longitude of ascending node, and longitude of perihelion) are assumed zero.

Planet	Semimajor axis (a)	Eccentricity (e)	True Anomaly (f)
Mercury	0.387098	0.20563	0°
Venus	0.723332	0.006772	180°
Earth	1.0000	0.0167	0°
Mars	1.5237	0.0934	180°
Jupiter	5.2044	0.0489	0°
Saturn	9.582	0.0565	180°

the orbit is related to M by noticing that the triangle ΔLCP has angles $C\hat{L}P = 180° - M'$, and $L\hat{P}C = M' - E'$. The side CP has length equal to a, and the side LC has length equal to ae. The students should find geometrically the relationship between eccentric and mean anomaly in this model. Applying the law of sines:

$$\frac{\sin(M' - E')}{ae} = \frac{\sin M'}{a} \tag{6}$$

that is,

$$\sin(M' - E') = e \sin M' \tag{7}$$

Thus, solving for E

$$E' = M' - \sin^{-1}(e \sin M') \tag{8}$$

At any time, the position of Mars, seen from the Sun, is again given by Eqs (4) and (5), except that now $E'(t)$ is non-uniform, given by Eq. (8). Again, the students should graph this model in comparison to "real" Mars. The result is shown in Figure 4. The students visualize that agreement is achieved to within half a degree. At this point the instructor should bring again to the students' focal awareness the critical aspect that Ptolemy did not have accuracy under a degree (Høg, 2017); thus, the students discern that the equant gives *excellent* agreement to the observations.

An extra assignment could consist of plotting the results of the equant model to the other planets visible in the pre-telescope era, as shown in Figure 5. This figure was computed with an equant for each planet. The "real" planets (Mercury, Venus, Earth, Mars, Jupiter and Saturn) were initialized as shown in Table 2, supposed massless, and alternating perihelion and aphelion.

Each planet has its own equant – which simply reflects the eccentricity of the orbit. Even in the case of Mercury, the planet of highest eccentricity, the agreement with the observations is satisfactory to the degree. The students should also conclude from this exercise how appropriate Mars was as subject of Kepler's analysis. The critical aspect now is that Tycho's observational data was accurate to 2 arcminutes (Høg, 2017). The instructor can ask: "based on this figure you produced, and knowing that Tycho's observations were accurate to two arcminutes, can you tell why Mars was appropriate for elucidating Kepler's laws?" The students should discern that, of the superior planets, Mars is the one whose deviation more blatantly disagreed with the prevailing model. Venus, Jupiter, and Saturn deviate by less than 2 arcminutes, within the accuracy of Tycho's data. Mercury, never too far from the Sun, is simply too difficult to observe.

4.5 Optional

The instructor may want at this point to do a parenthetical comment, returning for a moment to modern scientific parlance, and noting that the equant model is a model accurate to first order in eccentricity (Hoyle, 1973; Murray and Dermott, 1999). At higher eccentricities the equant model will again start to deviate significantly from the observations. Figure 6 shows a hypothetical planet of eccentricity $e = 0.45$. The equant model is off by more than 10 degrees. While for $e \sim 0.2$, like Mercury's orbit, the model is satisfactory down to 40 arcmin accuracy, for $e = 0.45$ one would need higher order corrections. Kepler's 2nd law is the full solution.

4.6 The ellipse has no equant

Ptolemy's solution had a very practical function. Given a point in the line of apsides upon which the planet sweeps equal angles in equal times, the orbit can be parametrized, as given by Eqs. (4) and (5) with E given by Eq. (8). Kepler had two problems. First, Tycho's observations of Mars, accurate to 2 arcmins, did not allow for the 30 arcmin error given by the equant model. Second, his ellipses, with the planet speeding nearing perihelion and slowing down nearing aphelion begged the question: *what is the equivalent to the equant*? What is the point about which a planet sweeps equal angles in equal times? Where is the point along the line of apsides that we can say that angle equals time? As it turns out, Kepler's quest to answer this question culminated with his 2nd law, that demolished the idea of uniform motion. The answer is: *there is no equant*. For the ellipse, there is no point about which an observer will see equal angles at equal times. *Kepler's first law does away with the epicycle. Kepler's second law does away with the equant*. This subsection is the one directly from Kepler's *Astronomy Nova*. I taught it in detail twice, before realizing that the level of detail is unnecessary. What the students should know is that **Kepler tried to find the equant for the ellipse and failed to find it**. In trying to find out the location of Mars' equant, Kepler again made use of Tycho's data. He took four observations of Mars in opposition, at times $t_1, t_2, t_3, and t_4$, corresponding to mean anomalies $M_1, M_2 M_3, and M_4$. The equant would be the unique point on the line of apsides whence Mars is seen at these angles (Fig. 7).

Yet, Kepler's best fit with an "ellipse equant" model was incompatible with the observations by 8 arcminutes, which was inadmissible by Tycho's 2 arcminute accuracy. Kepler had to go back and question his assumptions. But the assumptions were minimal. They amounted to:

1. Mars orbits the Sun;
2. Tycho's observations are reliable;
3. The equant exists.

(1) and (2) were beyond doubt correct. The conclusion was astonishing. The equant, a staple of astronomy for 1500 years, cannot exist. Kepler started this analysis by asking the question: where is the equant? And the answer was: there is no equant. There is no point about which we can say the planet sweeps equal angles at equal times. Uniform motion does not exist. Time is not given by angle. This is the aspect of the second law that should be emphasized: it rules out 1500 years of the paradigm of uniform motion.

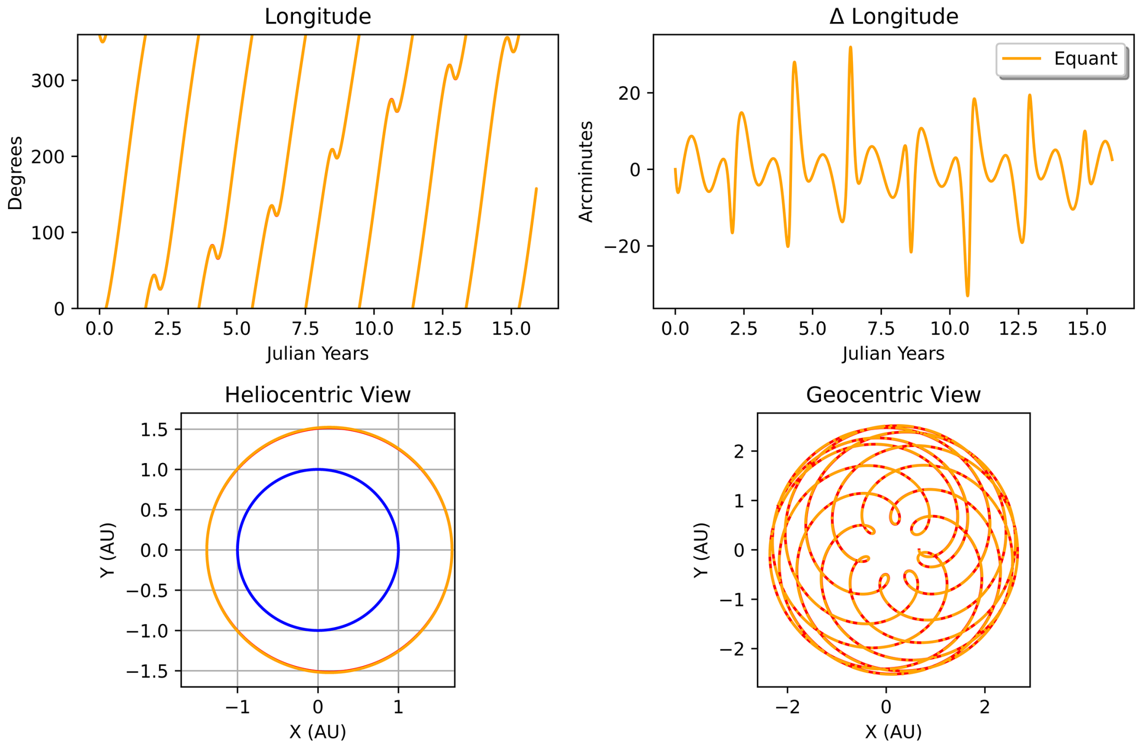

Figure 4. Mars orbit versus off-centered circular orbit model, with uniform motion about the equant. Upper left plot: Ecliptic longitude vs time, Mars (red) vs equant model (orange). Upper right plot: Longitude residual. The error is at most 30 arc minutes. Ptolemy's accuracy was 1 degree. The model is acceptable. Lower left plot: Heliocentric view of the orbit. Red is mars, blue is Earth, orange the equant model. Lower right plot: Geocentric view of the orbit. The equant model reproduces location and time of retrogradations.

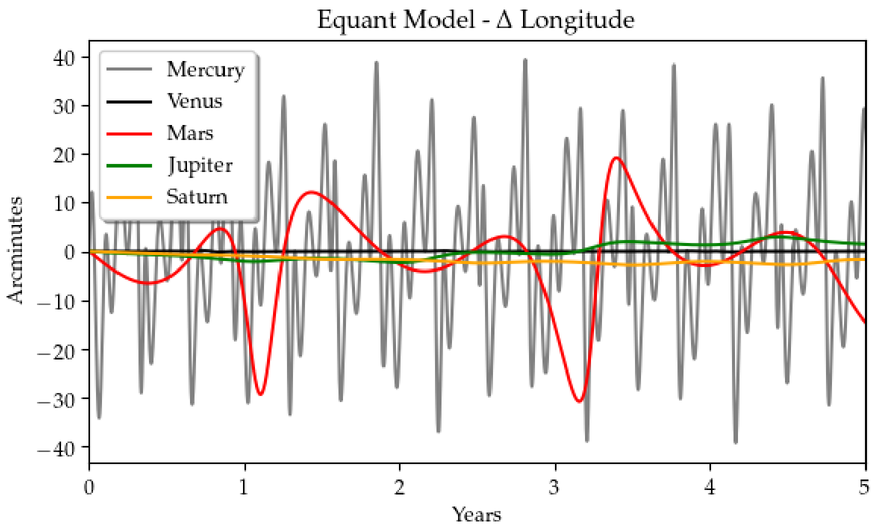

Figure 5. Residuals of the equant model for each planet. The residuals reflect orbital eccentricity. Even for Mercury, the most eccentric planet (e = 0.2), the equant agrees with the observations down to 40 arcminutes.

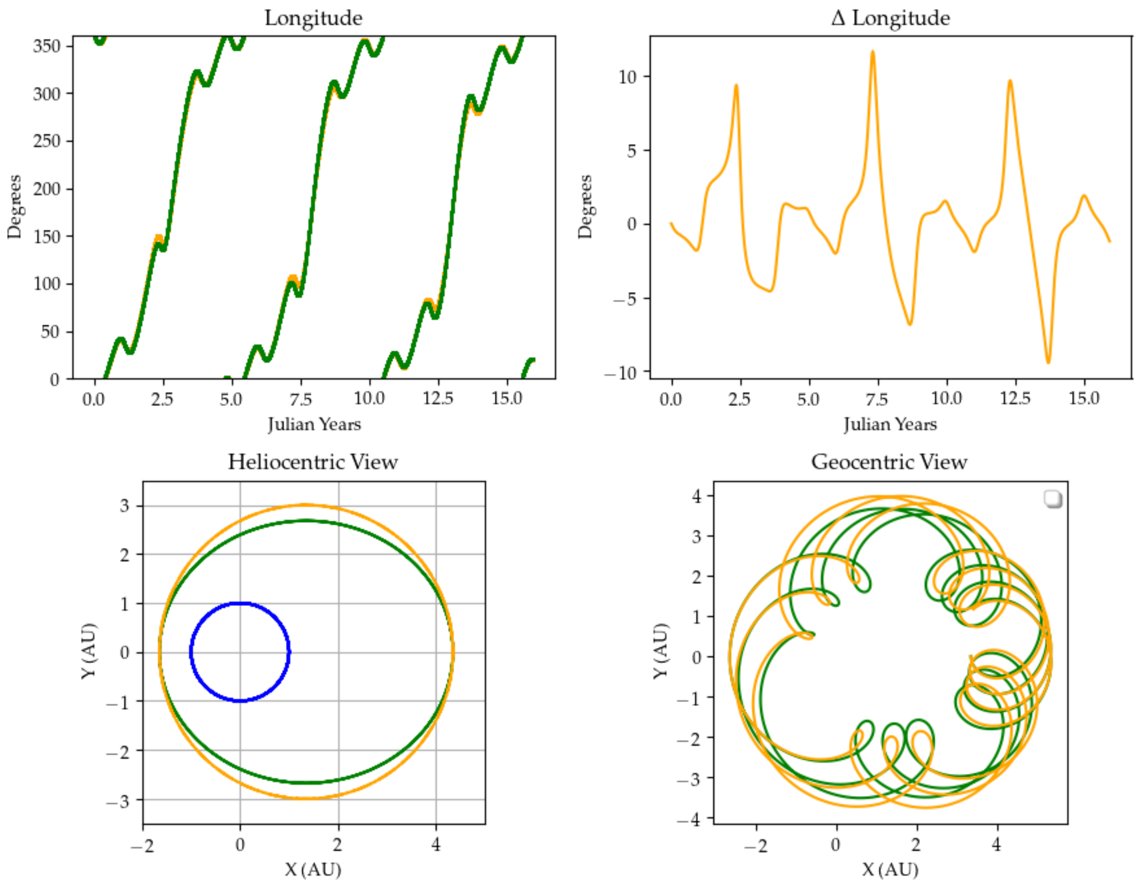

Figure 6. Validity of the equant model. A hypothetical planet (green) with orbital eccentricity 0.45, The equant model (orange) does not reproduce it well anymore. The equant is a model accurate only to 1st order in eccentricity.

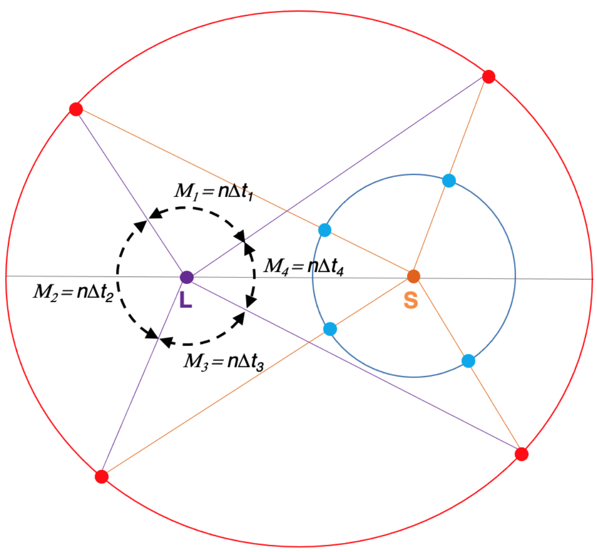

Figure 7. Kepler's method to determine the location of Mars' equant in an elliptic orbit, using four observations of Mars in opposition. The four dots in each orbit represent the four observations; the motion is counterclockwise. Because the Sun (S), Earth (blue orbit and dots), and Mars (red orbit and dots) are aligned, the orange lines intersect at the Sun. Timing the observations yields the mean anomalies M_1, M_2, M_3, M_4. Kepler then looked for the point L in the line of apsides where Mars would be seen at exactly these angles, failing to find it. The model could not be reconciled with the observations by 8 arcminutes, inadmissible by Tycho's 2 arcminute accuracy, forcing him to discard the equant model. The eccentricity of Mars' orbit is highly exagerated ($\sim$ 0.4 instead of $\sim$ 0.1) for clarity.

4.7 Time is equal area

Kepler had disproved fifteen centuries of "equal angles at equal times". That still leaves the problem of how to find the eccentric anomaly as a function of time. In looking for something that could be a measurement of time, Kepler stumbled on what this something was.

Because planets are slower when far from the Sun and faster when close, Kepler reasoned that the velocity was inversely proportional to the distance, $u \propto \frac{1}{r}$. If that is the case, one can multiply both sides by time and write the proportionality:

$$u\,r\,t \propto t \tag{9}$$

The quantity in the left-hand side, whatever it is, is linearly proportional to time. It is the "something" sought in equal "something" at equal times. But what is its interpretation?

The product ut is the length of the arc swung by the planet. If the time is infinitesimal, $t \rightarrow dt$, the arc is infinitesimal, $dl = udt$. Then the Sun, the planet's position at t and its position at $t + dt$ form a triangle, of area $r\,dl/2$. Comparing to Eq. (9), the quantity $u\,r\,dt = r\,dl$ that is linearly proportional to time is thus the area. As a planet orbits the Sun, the area it sweeps is proportional to time. Mean anomaly is not given by an angle. Mean anomaly is given by an area.

4.8 Kepler's Equation

At this point the class becomes finding the mathematical statement of Kepler's equation. The critical aspect to focus on is that the 2nd law is *quantitative*, relating the eccentric anomaly of the planet (and consequently its true anomaly) to the mean anomaly. The students have already studied in the 1st law the geometry of the ellipse (Appendix A).

The mean anomaly is proportional to the area. The question then is, how to compute the area? Kepler did not know calculus, so he could not calculate the area by summing the infinite infinitesimal distances. But Kepler was an excellent geometer. After he discovered that the orbit was an ellipse, he used the geometry of the ellipse to find out the area. Most students do not know a proof that the area of the ellipse of semimajor axis a and semiminor axis b is πab, so a short proof of it using integration is shown in Appendix B.

Given that the planet sweeps equal areas at equal times, a relation between mean anomaly (time) and area can be established. A planet sweeping equal areas at equal times will, within a time t, sweep an area

$$A_{sector} = \pi ab \frac{t}{T} \tag{10}$$

where T is the orbital period. The question is, what is the area of the sector? Consider Figure 8. The area swept from perihelion (X) to point P is the area of the elliptic sector A_{SPX}. We can relate it to the area of the circular sector A_{SQX} by using the relation $PH/QH = b/a$.

$$A_{SPX} = \frac{b}{a} A_{SQX} \tag{11}$$

The elliptic sector SQX can be broken down as the circular sector CQX, minus the triangle ΔCSQ

$$A_{SQX} = A_{CQX} - A_{CSQ} \tag{12}$$

The circular sector CQX comprises an angle E of the full 2π circle, so its area is

$$A_{CQX} = \pi a^2 \frac{E}{2\pi} \tag{13}$$

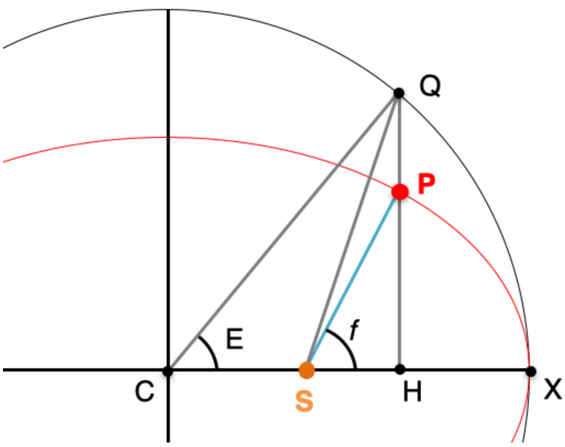

Figure 8. The true anomaly f is the angle, with vertex at the Sun, from perihelion to the planet. The eccentric anomaly E is the angle with vertex at the center of the orbit, from perihelion to the planet. Given the geometry of the ellipse, $PH/QH = b/a$, where b is the semiminor axis and a is the semimajor axis.

As for the triangle ΔCSQ, its base is $CS = ae$, and height $QH = a \sin E$

$$A_{CSQ} = \frac{1}{2} (ae) (a \sin E) \tag{14}$$

We thus find the area of the elliptic sector SPX,

$$A_{SPX} = \frac{ab}{2} (E - e \sin E) \tag{15}$$

But because area equal time, $A_{SPX} = \frac{\pi abt}{T}$. Equating both,

$$2\pi \frac{t}{T} = E - e \sin E \tag{16}$$

Given $n = \frac{2\pi}{T}$, the left hand side is $M = nt$, the mean anomaly. Thus,

$$M = E - e \sin E \tag{17}$$

This result is known as **Kepler's equation**. It is a direct consequence of Kepler's 2nd law, and can also be seen as Kepler's 2nd law itself. The left-hand side is time. The right-hand side is area. After teaching Kepler's 2nd this way, I gave as homework assignment a computational exercise where students had to predict the timing and location of Mars' next opposition, given the orbital elements and the current position of Mars and the Sun. I tested the method in both traditional and flipped classroom (King, 1993) environments. In the latter, the lab was started in an online class (taught during the 2020 pandemic of COVID-19), with the students sharing a *python jupyter notebook* via simultaneous video conferencing while freely debating. The method is thus in line with active learning (Bonwell and Eison, 1991), constructivist learning theory (Simon, 1995), and structured in a multimodal framework, featuring: videos, which are optimal for the visual/audio learners; notes, aimed at the reader learners; team-based learning, which helps the social learners; and coding exercises for the logical/math learners.

5 Assessment of Student Learning

A pre-and post- test was used in the graduate class administrated in 2019; the class had 10 students. The pre-test questionnaire had questions pertinent to all material covered in Fundamental Astronomy, among which one of the questions was "What is the relevance of area in Kepler's 2nd law?", to be answered discursively. The post-test was administrated after both Kepler's Laws and Celestial Mechanics modules, 3 weeks after Kepler's 2nd law was taught.

All students had the qualitative understanding of the law, as expected from graduate students in a major astronomy research institution, yet no one tied it to the crucial fact of finding the true anomaly. Post-instruction, the answers varied little from the pre-test, with 70% of the students repeating the "equal areas equal times" response of the pre-test. 30% showed a different answer: that area "allows position of the planet to be determined", "is time", and "shows there is no reference point from which an object appears to orbit equal angles in equal times". Perhaps the question of the pre-and-post test could have been better phrased. In contrast, anecdotal feedback I received so far, as well as the degree of student comprehension on dealing with Kepler's equation, supports the approach. One student in particular approached me to say they found the lab "inspiring" and " really interesting", adding in particular that they liked the way the class was taught, going through the historical development of the ideas and the critical thinking involved. One student, after the Mars opposition computational exercise, stated that “now I finally understand Kepler’s 2nd law and how to use it”. Another student said they liked how the class was taught “like research”. Yet another student declared “mad respect” for Kepler after the module.

6 Conclusion

In this work I constructed a way to teach Kepler’s 2nd law based on the historical method. The perceived benefits of the approach are enumerated below.

1. It frames the teaching in terms of a conceptual conflict, which research in human cognition shows is conducive to more effective learning (Piaget, 1977; Waxer and Morton, 2012). The conflict is between the equant model (equal angles) and Kepler’s 2nd law (equal areas).
2. Variational learning (Marton and Booth, 1997) is naturally brought to the classroom. In a scientific experiment, the impact of a variable is isolated by controlling its change while holding all the other variables constant, and analyzing the effect of the change. According to variational theory, this parallels how students construct learning. In the case in question, students explore different models until a match between data and model is achieved to observational accuracy. The different models hold one aspect constant while varying others (Sun centered/off-centered, and uniform/non-uniform motion).
3. The method uses the correspondence principle (Bohr, 1920), making the students understand the validity and limits of the equant model in its own historical framework, as a valid model that reproduces the observations up to about half a degree. The students understand why it worked (a model accurate to first order in eccentricity) and why it had to be discarded (when observational accuracy became better than the 40 arcminute accuracy given by the model). Data of worse quality would not have been able to discern between the equant model and Kepler’s 2nd law. As such, the approach also emphasizes to the students the paramount importance of observational accuracy.
4. The proposed approach highlights the revolutionary character of Kepler’s 2nd law: instead of repeating "equal areas in equal times" instructors can instead say "contrary to 1500 years of astronomical lore, there is no such thing as equal angles in equal times. Kepler’s 1st law discards the epicycle. Kepler’s 2nd law discards the equant. Area is how you measure time and hence how you find the planet."
5. The approach is also rooted in mathematical grounding, as “time = area” is stated not only through geometrical illustrations, but by Kepler’s equation (Eq. 17). Students can then manipulate it quantitatively, as usually done for the 3rd law.
6. By recreating the atmosphere of discovery, the method also frames the class in terms of cultural teaching (Matthews, 1989; Galili, 2010), bringing into the classroom the culture of astronomy. It is a narrative method that reveals the inner workings of the minds of the pioneers of the discipline, allowing their own voice to be brought into the classroom. For instance, Kepler famously wrote, in trying to locate the equant (Fig 7): “If this wearisome method has filled you with loathing, it should more properly fill you with compassion for me as I have gone through it at least seventy times at the expense of a great deal of time." That is a feeling that many a graduate student can empathize with. Combining human reason and emotion with the timeless elements of paradigm-shifting research, namely, tension between theory and new data, new data leading to a new theory, the new theory corresponding to the previous theory in its limit of applicability, this is a method that humanizes science and creates a bridge between a student experience and that of the greatest names in the history of the field.

Finally, on the limitations of the method, it has been said that the inquiry method, of which the historical method is a subset, is too difficult for any but the brightest students and that by teaching discarded ideas it is prone to causing confusion (Welch et al., 1981). Indeed, the method has been tested in upper division undergraduate and graduate studies only, so its effectiveness at lower division or general education courses is unconstrained. Also, because Kepler’s equation is transcendental, this method is best used in graduate curricula, where computational techniques are more routinely applied. Another criticism is that scientists are not historians and, by attempting to teach history of science we incur into the danger of teaching bad history (Matthews, 1989). Also, as stated by (Ausubel, 1968) “the most important factor influencing learning is what the learner already knows”. Indeed, resilience of previous conception is observed, as 70% of students repeated the pre- and post-test answer. The post-test was given after 3 weeks of instruction and after teaching Celestial Mechanics. It is unclear if the time elapsed, the introduction of Newtonian physics, or the phrasing of the question influenced this result.

7 Conclusion

I would like to acknowledge productive conversations with Joshua Tan, Mordecai Mac Low, and Roberto Pimentel. A first draft of this paper greatly benefited from comments by Anna Danielsson.

8 Author Biography

Wladimir Lyra is an Assistant Professor of Astronomy at New Mexico State University whose scholarship focuses on planet formation.

Appendix A: Elliptical orbits (Kepler’s 1st law)

In this appendix, the instructor guides the demonstration that the shape of the orbit is an ellipse. It works as active learning, with the instructor giving to each student a sheet with the elliptic shape, and a set of drafting tools. This appendix is written in teacher’s voice, guiding how to draw Figure 9 step by step. The order: is AP, circle centered at A, PH, circle centered at B, PQ, BQ, $\beta = H\hat{B}Q$, QK, AK, $AKQR$, and finally $\beta = A\hat{B}R$. A video is available at https://www.youtube.com/watch?v=Ss-nmWFY5Wo&list=PLatuGW739E01VsAwwqTHKU0tluD9h4c3I&index=3.

Having found the orbit, Kepler had no idea what geometrical shape it corresponded to. Yet, Kepler realized, through geometry, some properties this shape had. Consider Figure 9 (at this point with the only the red curve drawn, points B, C, P, major and minor axes; the focus A can be pre-drawn, or it can be found with the compass, from point F and striking the line of apsides with length BC, the semimajor axis). The Sun is at point A, and Mars at point P. The segment AP, of length r, is the radius vector from the Sun to the planet. The orbit is the red curve, which, a priori, we do not know what shape it corresponds to.

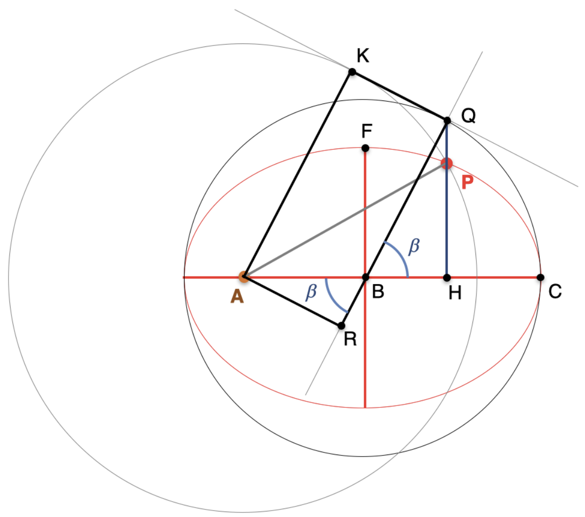

Figure 9. Kepler's method to find the curve corresponding to Mars' orbit. The crucial insight was to realize that a perpendicular to BQ at Q is tangent to the circle centered at the Sun (A) and with radius *r* equal to the radius vector that joins the Sun and the planet (P). The eccentricity of Mars' orbit is highly exagerated for clarity.

Some elements are:

1. The aphelion is point *C*;
2. The line of apsides is bisected at point *B*, the geometric center of the curve;
3. The length of the segment *BC* is by definition *a*;
4. The length of the segment *AB* is by definition *ae*; *e* is the eccentricity of the ellipse, but we do not know that yet. So far *e* is an *adhoc* constant, the factor by which we need to move the Sun away from the center.

What we want to find is *AP*, the radius vector of the orbit (draw AP). Algebraically, one would call it *r* and try to find a mathematical relationship for it. Geometrically, one draws a circle. This circle is centered at *A* and has radius *AP* (draw circle). We do not need to find exactly *AP*; if we find the radius of this circle, at any angle, we find the length of the radius vector. We trace the circle in the hope that a geometric coincidence that helps tell what the length *r* is becomes obvious.

Another way to find the shape of the curve is to find the coordinates *x* and *y* of the planet, and uncover their mathematical relation. We find the coordinate x by drawing the perpendicular from *P* to the line of apsides, defining point *H* (draw *PH* and point *H*). The coordinates of *P* are $x = BH$ and $y = PH$.

Next, we define the eccentric anomaly. For that we draw the circumscribed circle, of center *B* and radius $BC = a$ (draw circle). We prolong the line *PH* until it intersects the circumscribed circle at point *Q* (draw *PQ*). The eccentric anomaly is $E = A\hat{B}Q$ (draw *BQ*). We will also define the auxiliary angle $\beta = E - 180° = H\hat{B}Q$, (draw β). Given the triangle ΔBHQ, the coordinate *x* is $BQ \cos\beta$. Given $BQ = a$, we found the first coordinate.

$$x = a \cos\beta \tag{18}$$

As for *y*, the triangle ΔAHP can be used. It is a right triangle where $AP = r$ is the hypotenuse; the catheti are $PH = y$, and $AH = AB + BH = ae + x$. Thus,

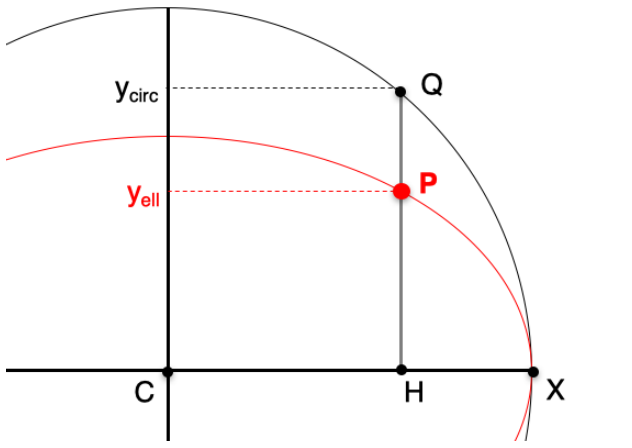

Figure 10. The ratio $PH/QH = y_{ell}/y_{circ}$ is equal to *b*/*a*, where *b* is the semiminor axis and a the semimajor axis of the ellipse.

$$y^2 = r^2 - (ae + x)^2 \tag{19}$$

Eq 18 gives the value of *x*, but the value of the radius vector *r* is so far unknown. Kepler found r in an ingenious way. He realized something curious: the perpendicular to *BQ* at *Q* is tangent to the circle of center *A* and radius *r*. (prolong *BQ* and draw the perpendicular)

Let *K* be the tangential point (draw *K*). Since *AK* is a radius (draw *AK*), then $AP = AK = r$. So, if we find *AK*, we find the value of *r*. Because *QK* is tangent to the circle, $A\hat{K}Q$ is a right angle. Kepler then prolonged the radius *BQ* to construct the rectangle *AKQR* (draw the rectangle and define *R*). Because this is a rectangle, $AK = QR = r$. The length *QR* is the sum of the radius ($BQ = a$) and the length *BR*. This length is given by the right triangle ΔARB. The hypothenuse is $AB = ae$, and the cathetus $BR = ae \cos\beta$ (draw $A\hat{B}R = \beta$). We find thus $AP = BQ + BR$, or $r = a + ae\cos\beta$. Given $\beta = E - 180°$.

$$r = a(1 - e\cos E) \tag{20}$$

Having found the radius vector, we substitute Eq. 20 and Eq. 18 into Eq. 19, finding

$$y^2 = a^2(1 - e^2)\sin^2 E \tag{21}$$

We can then write $\cos^2 E = x^2/a^2$, and $\sin^2 E = y^2/[a^2(1-e^2)]$ and invoke the trigonometric equality $\sin^2 E + \cos^2 E = 1$ to find the relationship between the coordinates

$$\frac{x^2}{a^2} + \frac{y^2}{a^2(1-e^2)} = 1 \tag{22}$$

This is the equation of an ellipse. The semimajor axis is a, and the semiminor axis is $b = a(1 - e^2)^{1/2}$.

Appendix B : Short proof of the area of the ellipse

Consider Figure 10. The main insight is that $PH/QH = b/a$, which is seen because $CQ = a$, and thus $QH = a\sin E$, and we have already proven (Eq 21) that $PH = a(1-e^2)^{1/2}\sin E = b\sin E$.

The area A_C of the circle is 4 times the area of the quadrant. The area of the quadrant can be found by integrating the vertical distances y_C from $x = 0$ to $x = a$.

$$A_C = 4 \int_0^a y_C \, dx \tag{23}$$

Given $y_C = a \sin E$ and $x = a \cos E$, then $A_C = \pi a^2$, as expected. The area of the ellipse is

$$A_e = 4 \int_0^a y_e \, dx \tag{24}$$

Because we can write $y_e = b/a \, y_C$, then

$$A_e = 4\frac{b}{a} \int_0^a y_C \, dx = \frac{b}{a} A_C = \pi ab \tag{25}$$

References

Aiton, E. J. (1969). Kepler's second law of planetary motion. *Isis*, 60(1):75–90.

Aktan, D. Ç. and Dinçer, E. O. (2014). examination of pre-service science teachers' understanding levels of kepler's laws with ranking task questions. *Journal of Baltic Science Education*, 13(2):276. http://oaji.net/articles/2015/987-1437063463.pdf.

Ausubel, D. (1968). *Educational psychology: a cognitive view.* New York: Holt, Rinehart & Winston.

Boccaletti, D. (2001). From the epicycles of the greeks to kepler's ellipse - the breakdown of the circle paradigm. In *Cosmology Through Time Ancient and Modern Cosmology in the Mediterranean Area*, Monte Porzio Catone (Rome), Italy, June 18-20, 2001.

Bohr, N. (1920). Über die serienspektra der elemente (on the spectral series of the elements). *Zeitschrift für Physik*, 2(5):423–469. https://doi.org/10.1007/BF01329978.

Bonwell, C. C. and Eison, J. A. (1991). *Active Learning: Creating Excitement in the Classroom. 1991 ASHE-ERIC Higher Education Reports*. ERIC. https://files.eric.ed.gov/fulltext/ED336049.pdf.

Brandenburg, A., Johansen, A., Bourdin, P. A., Dobler, W., Lyra, W., Rheinhardt, M., Bingert, S., Haugen, N. E. L., Mee, A., Gent, F., et al. (2021). The pencil code, a modular mpi code for partial differential equations and particles: multipurpose and multiuser-maintained. *Journal of Open Source Software*, 6(58):2807.

Bussey, T. J., Orgill, M., and Crippen, K. J. (2013). Variation theory: A theory of learning and a useful theoretical framework for chemical education research. *Chemistry Education Research and Practice*, 14(1):9–22.

Carroll, B. W. and Ostlie, D. A. (2007). *An introduction to modern astrophysics and cosmology*. Pearson, San Francisco.

Chambers, J. E. (2012). Mercury: A software package for orbital dynamics. *Astrophysics Source Code Library*, pages ascl–1201.

Cheng, E. W. (2016). Learning through the variation theory: A case study. *International Journal of Teaching and Learning in Higher Education*, 28(2):283–292.

Coelho, R. L. (2009). On the concept of energy: how understanding its history can improve physics teaching. *Science & Education*, 18(8):961–983. https://doi.org/10.1007/s11191-007-9128-0.

Conant, J. B. (1964). *Case studies in experimental science*. Harvard university press.

Copernicus, N. (1543). *On the Revolution of the Heavenly Spheres*.

de Cayeux, A. and Brunier, S. (1983). Les planètes. *Ciel et Terre*, 99:29.

Dixon, F. A., Prater, K. A., Vine, H. M., Wark, M. J., Williams, T., Hanchon, T., and Shobe, C. (2004). Teaching to their thinking: A strategy to meet the critical-thinking needs of gifted students. *Journal for the Education of the Gifted*, 28(1):56–76. https://doi.org/10.1177/016235320402800104.

Duhem, P. (1954/1906). *The aim and structure of physical theory*. Princeton University Press, Princeton: NJ.

Galili, I. (2010). History of physics as a tool for teaching. In Vicentini, M. and Sassi, E., editors, *Connecting research in Physics Education with Teachers Education.*, pages 153–166. I.C.P.E. book.

Gingerich, O. (1983). Laboratory exercises in astronomy–the orbit of mars. *Sky and Telescope*, 66:300.

Guimaraes, L. A. and Fonte Boa, M. (2006). *Fisica Volume 1: Mechanica. 3rd Edition*. Hipermidia Guys, Brazil.

Halliday, D. and Resnick, R. (2020). *Physics for Students of Science and Engineering*. John Wiley & Sons Inc.

Høg, E. (2017). Astrometric accuracy during the past 2000 years. *arXiv preprint arXiv:1707.01020*. https://arxiv.org/abs/1707.01020.

Holton, G. (1978). On the educational philosophy of the project physics course. In *The Scientific Imagination: Case Studies*. Cambridge University Press, Cambridge.

Holton, G. and Brush, S. (1952). *Introduction to concepts and theories in physical science. 3rd edition: Holton, G., Brush, S. & Evans, J. (2001). Physics, the Human Adventure: From Copernicus to Einstein and Beyond*. Rutgers U. P., New Brunswick, N.J.

Hoyle, F. (1973). *Nicolaus Copernicus: an essay on his life and work. 1st edition*. Heinemann., London.

Kepler, J. (1609). *Astronomia Nova*.

King, A. (1993). From sage on the stage to guide on the side. *College teaching*, 41(1):30–35. https://doi.org/10.1080/87567555.1993.9926781.

Kuhn, T. (1962). *The structure of scientific revolutions*. University of Chicago Press, Chicago, IL.

Leinhardt, G. (1988). Getting to know: Tracing students' mathematical knowledge from intuition to competence. *Educational Psychologist*, 23(2):119–144. https://doi.org/10.1207/s15326985ep2302_4.

Mach, E. (1872/1911). *The history & root of the principle of the conservation of energy*. Open Court, Chicago, IL.

Mach, E. (1943/1895). On instruction in the classics & the sciences. In *Popular Scientific Lectures*. Open Court. La Salle, Chicago.

Mansouri, F. A., Tanaka, K., and Buckley, M. J. (2009). Conflict-induced behavioural adjustment: a clue to the executive functions of the prefrontal cortex. *Nature Reviews Neuroscience*, 10(2):141–152. https://doi.org/10.1038/nrn2538.

Marion, J. and Thornton, S. (2020). *Classical Dynamics of Particles and Systems. 3rd Edition*. Harcourt Brace Jovanovich, New York, NY.

Marton, F. and Booth, S. (1997). *Learning and Awareness*. Lawrence Erlbaum, Mahwah, NJ.

Marton, F. and Pang, M. F. (2008). The idea of phenomenography and the pedagogy for conceptual change. In Vosniadou, S., editor, *International handbook of research on conceptual change*, pages 553–559. Routledge, London, UK.

Marton, F., Runesson, U., and Tsui, A. B. M. (2004). The space of learning. In Marton, F., Tsui, A. B. M., Chik, P. P., Ko, P., and Lo, M., editors, *Classroom discourse and the space of learning*, pages 43–62. Lawrence Erlbaum, Mahwah, NJ.

Matthews, M. R. (1989). A role for history and philosophy in science teaching. *Interchange*, 20(2):3–15. https://doi.org/

10.1007/BF01807043.

Monk, M. and Osborne, J. (1997). Placing the history and philosophy of science on the curriculum: A model for the development of pedagogy. *Science education*, 81(4):405–424. https://doi.org/10.1002/(SICI)1098-237X(199707)81:4%3C405::AID-SCE3%3E3.0.CO;2-G.

Murray, C. D. and Dermott, S. F. (1999). *Solar System Dynamics*. Cambridge University Press.

Orgill, M. (2012). Variation theory. In Seel, N. M., editor, *Encyclopedia of the sciences of learning*, page 3391–3393. Springer-Verlag Gmbm, Heidelberg, Germany.

Pang, M. F. and Marton, F. (2003). Beyond "lesson study": Comparing two ways of facilitating the grasp of some economic concepts. *Instructional Science*, 31(3):175–194. https://doi.org/10.1023/A:1023280619632.

Piaget, J. (1970). *Genetic Epistemology*. Columbia University Press, New York, NY.

Piaget, J. (1977). *The development of thought: Equilibration of cognitive structures*. Viking, New York, NY.

Posner, G. J., Strike, K. A., Hewson, P. W., et al. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science education*, 66(2):211–227. https://doi.org/10.1002/sce.3730660207.

Rein, H. and Liu, S.-F. (2012). Rebound: an open-source multipurpose n-body code for collisional dynamics. *Astronomy & Astrophysics*, 537:A128.

Rice, M. J. (1972). The case for the disciplines in the organization of social studies curricula for elementary and secondary schooling. paper presented at the colleges and university faculty association of the rational council for the social studies. Technical report, Boston, Massachussets.

Ryden, B. and Peterson, B. (2020). *Foundations of Astrophysics*. Cambridge University Press, Cambridge.

Saumure, K. and Given, L. (2008). Convenience sample. In Given, L., editor, *The SAGE encyclopedia of qualitative research methods*, pages 125–126. SAGE Publications, Thousand Oaks, CA.

Schwab, J. J. (1962). The teaching of science as inquiry. In Schwab, J. J. and Brandwein, P., editors, *The teaching of science*, pages 125–126. Harvard University Press, Cambridge, MA.

Setyadin, A. H., Suryana, T. G. S., Utari, S., Efendi, R., Liliawati, W., and Utama, J. A. (2020). Identifying k-10 students' learning difficulties on learning kepler's law using worksheet: Is it worth? In *Journal of Physics: Conference Series*, volume 1467, page 012051. IOP Publishing. https://doi.org/10.1088/1742-6596/1467/1/012051.

Simon, M. A. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for research in mathematics education*, 26(2):114–145. https://doi.org/10.5951/jresematheduc.26.2.0114.

Waxer, M. and Morton, J. B. (2012). Cognitive conflict and learning. In Seel, N. M., editor, *Encyclopedia of the Sciences of Learning*, pages 585–587. Springer US, Boston, MA. https://doi.org/10.1007/978-1-4419-1428-6_280.

Welch, W. W., Klopfer, L. E., Aikenhead, G. S., and Robinson, J. T. (1981). The role of inquiry in science education: Analysis and recommendations. *Science Education*, 65(1):33–50. https://onlinelibrary.wiley.com/doi/abs/10.1002/sce.3730650106.

Yu, K. C., Sahami, K., and Denn, G. (2010). Student ideas about kepler's laws and planetary orbital motions. *Astronomy Education Review*, 9(1). https://doi.org/10.3847/AER2009069.

Zazkis, R. and Chernoff, E. (2006). Cognitive conflict and its resolution via pivotal/bridging example. In Novotná, J., Moraová, H., Krátká, M., and Stehlíková, N., editors, *Proceedings of the 30th Annual Conference of the International Group for the Psychology of Mathematics Education 5*, pages 465–472, Prague. PME. http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.615.9518.

ASTRONOMY EDUCATION JOURNAL

Astronomy Education Journal, 2021, Vol. 01, No. 1

doi: 10.32374/AEJ.2021.1.1.005

RESOURCES & ACTIVITIES

An observational project for a large class – determination of the duration of the sidereal day

William Tobin[1,*,†]

[1]Retired from: School of Physical and Chemical Sciences, University of Canterbury, Private Bag 4800, Christchurch 8140, New Zealand
Present address: 6 rue Saint Louis, 56000 Vannes, France

[*]william@tobin.fr
[†]ORCID: 0000-0002-0533-411X

Abstract

Having students confront the real night sky is difficult for a large class. The determination of the duration of the sidereal day is, however, a project that students can do on their own without specialized equipment. This project provides an introduction to how observational science is done because students must devise the observational procedure, make timings, analyse them, and present results including uncertainties in a report. The experience of running this project as part of a first-year university introduction to the grand ideas of physics is described, with suggestions for improvements and to reduce cheating. Almost a third of students reported results within 1 second of the accepted value.

Keywords: Sidereal day; Observational project; Large classes

I saw Eternity the other night,
Like a great ring of pure and endless light,
All calm, as it was bright;
And round beneath it, Time in hours, days, years,
Driv'n by the spheres
Like a vast shadow mov'd; in which the world
And all her train were hurl'd.

— From *The World*, by the Welsh metaphysical poet Henry Vaughan (1621-1695), used as an epigraph in a student report

1 Introduction

Observational work involving quantitative results presents problems for the teacher of introductory astronomy. Having students themselves make observations which they will subsequently analyse usually requires expensive and delicate equipment, long bouts of supervision outside normal teaching hours, and contingency arrangements in case of cloud. These difficulties are multiplied if the class is large. Yet an introductory astronomy class in which the student does not confront the real sky is hollow.

There are, however, some projects that students can undertake on their own with minimal equipment. Observations of the lengths of midday shadows is one (Jackson, 2004; Kwok, 2004) but requires student presence around noon, which can lead to timetable clashes in a higher-education context. Understanding variations of the lunar diameter, measured using a home-made sighting device (Krisciunas, 2010), is probably too complex – and the duration of observations too long – for an introductory class. In this note, I describe a short and conceptually-simple project to determine the duration of the sidereal day that I ran for six years (1988-1993) as part of the term-long 'observable universe' component of a 'grand ideas of physics' first-year course for arts and science students at the University of Canter-

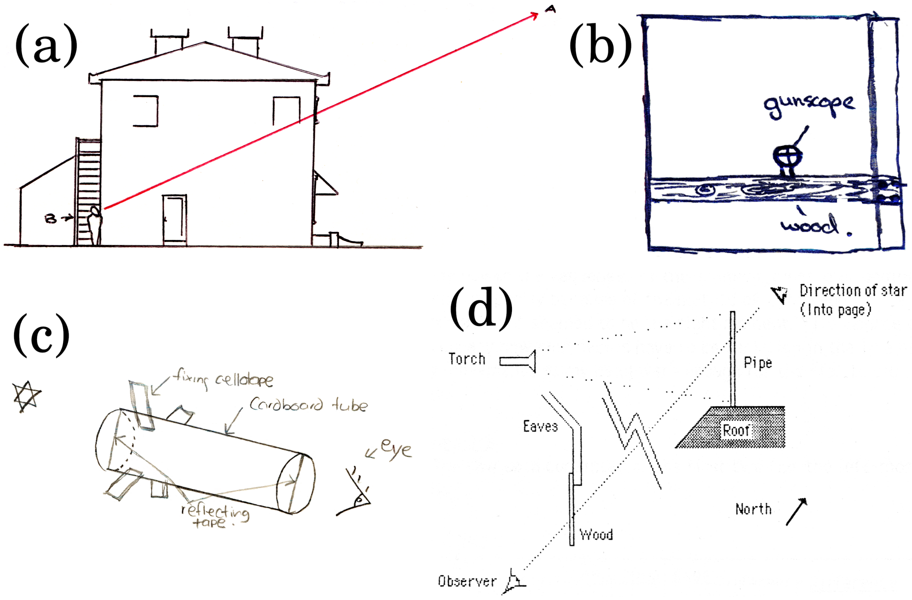

Figure 1. Various student sight lines. (a) The north-south side of a house. (b) A telescopic rifle sight lashed to a wooden bar nailed across a window frame. Considerable relashing was necessary to centre the target star (Rigel) in the gun sight. (c) The parallel edges of reflecting tape across the ends of a toilet roll, with the roll itself taped to a window. Though this had the advantage of an ~20° altitude range for chosing the target star, the short tube length produced inaccurate results. (d) Alignment between a wooden batten nailed to the eaves of the student's house and an overflow pipe many metres away on a neighbour's roof. The pipe needed illumination with a torch to be visible. Now that cameras are easily available in mobile phones, student reports will doubtless often include photographs of the sight line.

bury in Christchurch, New Zealand. The full course involved two lectures per week for three terms, or about 54 lectures in total, and no backround in science or mathematics was assumed beyond early secondary school. Class size was about one hundred, and from the instructor's point of view, the only effect of a larger or smaller number would be on the time required for marking.

For northern-hemisphere observers Monson (1992) suggests finding the sidereal-day duration by plotting the positions of Polaris and circumpolar stars over two or more hours using a plastic sheet taped against a north-facing window. 'Careful students can easily come within 15 minutes of the accepted value of $23^{h}56^{m}$,' he states. For those with a reflex camera, star trails provide an analogous method (Royal Observatory Greenwich, 2015). Eckroth (1996) describes using a fixed telescope and an equatorial star to determine a value close to the expected duration.

The approach outlined here is similar to Eckroth's, but requires no telescope. It aims to give students a taste of how quantitative science is done. Besides forcing the students to confront the sky, the project provides an introduction to the process of experimental science. The student must use initiative and pay attention to detail in planning, learn to recognize a brightish star, and make half-a-dozen or so timings. The project requires simple graphical analysis and the presentation of results and uncertainties in a report. I shall suggest how some of the problems encountered three decades ago can now be mitigated thanks to technological advances.

2 Instructions to Students

The sidereal-day duration is determined by timing the transit of a given star across a fixed sight line over several nights. Part of a lecture and a three-page handout provided guidance to students.

2.1 Star and sight line

Each student's first task was to identify a recognisable star and find or set up a suitable sight line that it will cross when he or she expects to be free from other obligations such as residence-hall dinner (or the pub). For greatest accuracy, the sight line should be on the meridian and the star on the celestial equator. This equatorial requirement was not suggested so as not to exclude sight lines to the south for which the stars of Crux and the pointers might be known to the students. Nevertheless, some students realized the desirability of an equatorial star. It was suggested that about half a dozen timings should be taken over an interval of about three weeks. In Aotearoa/New Zealand (the Land of the Long White Cloud), poor weather (and social commitments!) can easily extend this interval, against which eventuality students were asked to ensure that their sight line would still be crossed during the hours of darkness. Students were asked to identify their star on a star chart, not least to avoid any confusion with planets, which would yield incorrect results.

Students showed considerable ingenuity in finding their sight lines (Fig. 1). The corners of university buildings and a reproducible observation point provided many. North-south building walls furnished others. Chimney stacks, water towers and lamp posts were employed. In New Zealand many students live at home, and they enjoyed wider possibilities, such as stakes attached to fences or driven in the ground, or nails hammered into walls, or sighting through plastic pipes. In a land where hunting is common, several students were able to use telescopic rifle sights. The individual nature of each student's star and sight line was of course a brake impeding copying of other students' observations.

As a whole, students were not good at identifying their target stars by name. Choosing a star and sight line are not independent tasks and many target stars were doubtless not the brightest ones that appear on simple star charts. Stars that students did identify by name were bright ones: Regulus, Betelgeuse, Rigel, Alphard, Sirius, Canopus and α & β Crucis, though students noted that with a very bright star, like Sirius, its glare could make timing difficult. Nowadays, the identification problem is solved, as almost all students own orientation-sensitive smartphones. Star identification is easy with freeware versions of applications such as SkyView, which shows and identifies the stars towards which the 'phone is pointing, down to fainter than 4th magnitude.

Sight line and star selected, students could encounter unexpected difficulties, such as chalk marks that washed off in the rain, or rugby-club floodlights that rendered the target star invisible.

2.2 Timing

Students were asked to ensure that their timepiece was always set to accurate time. For students in the late 1980s/early 1990s that meant repeated comparisons with radio or TV time signals, or the then-new talking-clock service or teletext clock. Nowadays, accurate time is more easily available. Smartphones synchronize to standard time by a variety of means: from the mobile network, from GPS satellites or if connected to the internet via the Network Time Protocol (NTP), as do web sites such as `time.is`, `clock.zone` and `time.gov`. Students were asked to state in their reports how they ensured timings linked to standard time. Some missed the important point that their timepiece needed to be synchronized repeatedly because a watch that gains or loses a few seconds a day leads to a result incorrect by this amount. This is unlikely to be a problem now.

When making their timings, students needed to state their criterion for establishing the instant of crossing and the uncertainty they attached to it. Was one eye or two used? Averted vision? One student's sight plane was defined by a window frame and a distant lamp post, and refined by a needle with a large eye that was stuck into the window frame. The student's eye was positioned so that timings were made when the star filled the needle's eye. Depending on how careful they were, students estimated the uncertainty in a single timing to be from 5 to 30 seconds. Those using a telescope or gun sight estimated 1 or 2 seconds.

2.3 Analysis

The students were asked to plot their crossing times against date. The (negative) slope, in seconds per day, indicates how much shorter the sidereal day is than 24 hours. Students were not expected to know about linear regression (although the occasional one did), so unless a student had been very careless, this first, hand-plotted graph did not provide enough resolution to extract the full accuracy of the measurements: a pencil line would be thicker than the uncertainties in the timings. Students were therefore asked to round their derived slope to a whole number of seconds and to calculate for each observation (i) the crossing time predicted assuming this rounded slope, and (ii) by how much the observed crossing time differed from this prediction. Plotted against date, the slope of this second graph of differences provided a correction. In addition, this second graph allowed the exclusion of clearly discrepant observations. The timing uncertainty of the retained observations and their scatter were used to provided an estimate of the uncertainty in the final result via lines with a greater and a lesser slope that the students felt were still, just, compatible with their observations (Fig. 2).

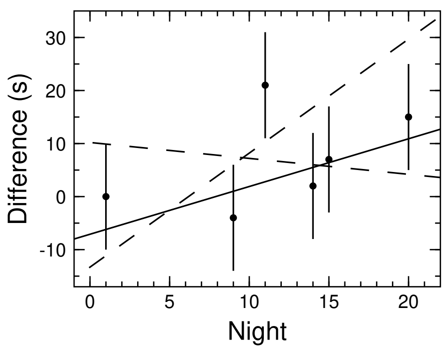

Figure 2. This student's crossing times plotted against night number yielded a whole-second sidereal-day duration of $23^h\ 56^m\ 03^s$. This second plot compares the crossing times against those expected using this first-estimate duration. The full line is the student's by-eye estimate of the best fit and has a slope of 0.9 s/day. The student (the same as in Fig. 1 (d)) estimated timings uncertain by 10 s, which from their scatter appears to be a slight over-estimate, and excluded the timing on night 11 which he felt indicated some blunder. The dashed lines show the extreme slopes the student considered just consistent with the timings, yielding a final estimate for the duration of the sidereal day of $23^h\ 56^m\ 03.9^{s+1.3}_{-1.2}$.

Table 1. Marking scheme.

Item	%
Introduction and explanation of the sidereal day	10
Description of the observational procedure, including orientation of the sight line, choice of star and timing	25
Presentation of data	10
Analysis of observations	15
Discussion of individual and final uncertainties	15
Suggested improvements	15
Clarity of exposition and analysis	5
English	5

Many students found this analysis difficult, and there were some suprising pitfalls. A few students were confused by the change from daylight-saving to standard time in those years when it occured during the project. Some less-numerate students thought that the interval between day 1 and day N was N, not $N - 1$. Another common error came in the calculation of expected crossing times. Students usually chose one of their actual timings to correspond to zero difference, but did not appreciate that this was a point just like any other in the second graph. Students did not always realize that expected times needed to be calculated from the round-number slope rather than the accepted value of the sidereal day – though in practice, of course, this is a self-correcting error. Further comments follow in Section 3.1.

Students were asked to write a report presenting their work, including suggestions for improvements. It was suggested this report did not need to be as long as 2000 words,

3 Student results

3.1 The duration of the sidereal day

Grading the first student reports in 1988 I immediately encountered two difficulties. First, it was necessary to devise a marking scheme to provide grading uniformity (Table 1). Secondly, checking the analysis in a hundred or so reports was impossibly time consuming. It was necessary to write an interactive, linear-regression computer program into which I could quickly input each student's observations in order to see if the claimed results were in accord with the reported timings. The program made it easy to exclude any outlier points.

Fig. 3 (a) shows the results and uncertainties claimed by students in the 1991 project year. As in earlier years, the majority of students produced results close to the accepted duration of $23^h\ 56^m\ 04.0905^s$.

A handful of results clustered near $23^h\ 55^m\ 50^s$. The cause

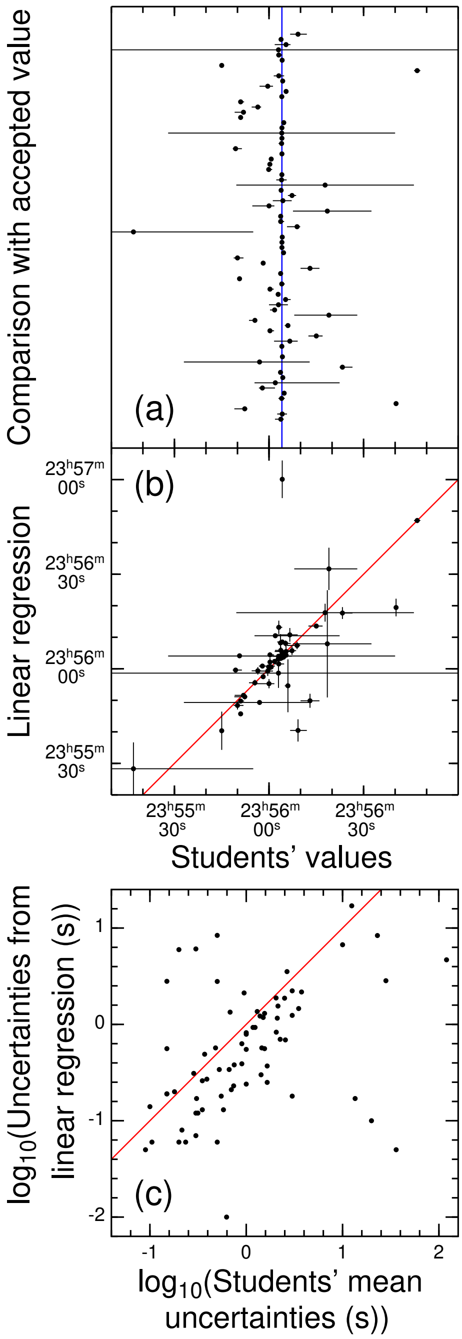

Figure 3. Results from 1991. Not plotted are one result far outside the box and four where no uncertainty was estimated. (a) Stack plot of the sidereal-day length and its uncertainty, as derived by the students. The claimed uncertainty is often smaller than the symbol size. The blue vertical line indicates the accepted value of $23^h\ 56^m\ 04.1^s$. Forty-seven out of 74 results fall within ±4 s of this value; 39 within 2 s; and 32 within 1 s. Students using gun sights produced results accurate to 0.1 s. There is a small cluster of results around $23^h\ 55^m\ 50^s$, discussed in Section 3.1. (b) The students' reported sidereal-day duration compared with my linear-regression analysis of their timings. The red diagonal line indicates equality. The values differ by less than ±4 s in 53 out of 74 cases. (c) It is difficult to compare the students' final uncertainties with those I derived from linear regression in panel (b), so here they are plotted on logarithmic scales, where the red diagonal line indicates equality.

was as follows. To illustrate the graphical analysis, invented timings were presented and analysed in the handout. They produced a sidereal-day duration close to $23^h\ 55^m\ 50^s$. Despite the fact that it was stated in bold face that these timings and the result were entirely fictitious, some students evidently thought their own result should accord with the example.

Did some students delude themselves, mistakenly expecting crossing times ~$4^m\ 10^s$ earlier each night? No doubt some did, but over three weeks that corresponds to an accumulated error of an improbable 5 minutes. A few students clearly got lost in the graphical analysis and sexagesimal arithmetic, with derived durations far from those implied by their reported timings (Fig. 3 (b)). However, a rump of students evidently falsified their timings, reporting ones that were highly consistent with the fictitious $23^h\ 55^m\ 50^s$ value given in the handout. Another fraud was clear when a student looked up the expected solar-sidereal difference correctly, but applied it in the wrong sense in his fabricated observations, producing a target star that crossed *later* each night. Perhaps falsified timings are unsurprising. Inventing and altering data were among the top cheating activities amongst English university students reported by Newstead et al. (1996).

To counter all this, I strengthened the handout wording and lecture presentation, suppressing words such as *fictitious*, which it transpired were unknown to some students. I also asked students to present their project report in two halves, the first with their timings, so they would be committed to them, and the second with their analysis. The number of results near $23^h\ 55^m\ 50^s$ declined from its peak in the first year but was not completely extinguished, as Fig. 3 (a) testifies. One must also wonder how many students were savvy enough to escape detection with fabricated timings based on the true sidereal-day duration.

Nowadays, with easy internet connectivity, it would be possible for the instructor to set up an interactive web page on which students would report their crossing times within, say, 48 hours of making them. This should reduce the prevalence of fraud. In addition, various checks could be programmed. Are observations reported for nights that were cloudy? If the target star is known, does the the crossing time correspond to the azimuth of the claimed sight line? If by-hand analysis is felt to be too much for the students, the analysis could be programmed into the web page, or effected with other tools such as spreadsheet software like Excel.

Figure 3 (c) compares the uncertainties stated by the students with those I obtained by linear regression. There are many cases where the students have significantly under- or over-estimated their uncertainties, but it is striking that most are essentially in agreement with the linear-regression values, although larger by a factor of about two (0.3 in the logarithm). The students were asked to estimate the *range* of sidereal-day durations just consistent with their data. The mean uncertainty plotted here is half this range and obviously this 'two-σ' value would be expected to be about twice that obtained by linear regression, as seen.

One year a number of students stated that the accepted length of the sidereal day was $23^h 56^m 03^s$, which on investigation was the erroneous value given by one of the astronomy text books in the University Library (Jastrow and Thompson, 1984). An occasional, well-informed student noted that because of the precession of the equinoxes the sidereal day is shorter than the rotation period of the Earth by 0.0084 s.

3.2 Suggested improvements

Students suggested a number of possible improvements, such as a head brace to keep the eye still, working in pairs with one student watching the star and the other the time, working in the less-cloudy summer months, working away from the city and its lights, using a watch with a lighted dial, using a fixed telescope or gun sight, observing over a longer interval – and restricting social life! One student invoked the possibility of photographing star trails. A history student with math anxiety suggested help from a more-numerate friend. 'Finally, if I were to repeat this experiment,' wrote another student, 'I would not take my readings outside a house that has a great big dog stationed on the front lawn.'

4 Conclusion

Asking students to determine the duration of the sidereal day is a project that is suited to large classes, and could even be used in on-line courses. Besides introducing students to the sky, it parallels real research – devising an observational procedure, making observations, analysing them, and reporting the results. These are of course useful skills beyond introductory astronomy. As with all independent work, there are opportunities for cheating but they can be reduced by using the strategies outlined.

Many students were thrilled by their interaction with the real night sky. 'This project was, in my opinion, very exciting and satisfying,' wrote one. As a university teacher, I have always thought, to paraphrase Plutarch, that education is the kindling of a flame, not the filling of a vessel (Babbitt, 1927). I was particularly touched by the words of another student: 'This project has made me stop and take stock,' she wrote. 'Previously I thought that a day was a day, 24 hours, now I have found another type. I wonder how many more things I take for granted have hidden or partly hidden companions!!'

References

Babbitt, F. C. (1927). *Plutarch's Moralia*, volume I. Harvard University Press, 1st edition. 'On Listening to Lectures,' at p. 257.

Eckroth, C. A. (1996). Observations of the sidereal day. *The Physics Teacher*, 34:94. https://doi.org/10.1119/1.2344355.

Jackson, E. (2004). Daytime astronomy in the northern hemisphere using shadows. *Astronomy Education Review*, 2:146–149. https://doi.org/10.3847/AER2003021.

Jastrow, R. and Thompson, M. H. (1984). *Astronomy: Fundamentals and Frontiers*. Wiley, 4th edition. p. I-26.

Krisciunas, K. (2010). Determining the eccentricity of the Moon's orbit without a telescope. *American Journal of Physics*, 78:834–838. https://doi.org/10.1119/1.3369968.

Kwok, P.-W. (2004). Studying the transit of the Sun using shadows. *Astronomy Education Review*, 3:106–114. https://doi.org/10.3847/AER2004010.

Monson, B. (1992). A simple method of measuring the length of the sidereal day. *The Physics Teacher*, 30:558–559. https://doi.org/10.1119/1.2343640.

Newstead, S. E., Franklyn-Stokes, A., and Armstead, P. (1996). Individual differences in student cheating. *Journal of Educational Psychology*, 88:229–241. https://doi.org/10.1037/0022-0663.88.2.229.

Royal Observatory Greenwich (2015). Using star trails to measure the sidereal day. Retrieved from https://www.rmg.co.uk/sites/default/files/KS4%20Star%20Trails%20and%20Sidereal%20Day.pdf (accessed 2021 January 14).

Astronomy Education Journal, 2021, Vol. 01, No. 1

doi: 10.32374/AEJ.2021.1.1.006

RESOURCES & ACTIVITIES

Timing pulsars: An exercise in statistical analysis and the scientific process

John R. Walkup[1,*,2], Joseph White[1] and Roger Key[1]

[1]Department of Physics, California State University, Fresno, Fresno, California, USA and [2]Present Affiliation: College of Mathematics and Science, University of Central Oklahoma, Edmond, Oklahoma, USA

[*]jwalkup2@uco.edu

Abstract

A lab activity for teaching students the fundamentals of statistical analysis by timing pulsar periods is described. The electromagnetic pulses of pulsars have been mapped to sound and uploaded on social media channels, allowing students to "listen" to the beat of the pulsar. Because these beats are extraordinarily precise, they can serve as cyclic events of known time duration. The three-step process described in this article first requires that students select a timing method of low random error found by comparing standard deviations between two suggested methods. In the second step, students reduce systematic error by calibrating their optimal method using a pulsar of known time duration. Finally, students time an unknown pulsar (the mystery pulsar) using the optimal method chosen in Step 1 and calibrating out the bias found in Step 2. By expressing their results in terms of confidence intervals, they use a professional pulsar database to identify the mystery pulsar. Because students are not informed of the identity of the pulsar until after they turn in their lab reports, they are compelled to perform the measurements as carefully and objectively as possible. This activity provides a perfect vehicle for astronomy labs at the beginning of a semester — including online instruction — because it requires no prior instruction in astronomy and no equipment other than the stopwatch on a cell phone and internet connection. Furthermore, this activity offers an introduction to pulsars and such physics topics as magnetism and the conservation of angular momentum.

Keywords: Pulsars; Error analysis; Standard deviation; Standard error; Confidence intervals

1 Introduction

1.1 Background

A recent article described a three-step procedure for teaching students to conduct scientific measurement that reduces random and systematic error (Walkup et al., 2019). That article relied on the use of a light pulse generated by an Arduino. The activity described here uses the same three-step procedure, but replaces the Arduino with the electromagnetic pulse of a pulsar.

The pedagogical purposes of the lab activity discussed here mirrors that in Walkup et al. (2019). For one, this activity attempts to supplant the traditional verification lab, where students simply complete stepwise procedures to verify physics concepts learned in lecture, with more career-oriented experimental strategies likely to be used by practicing scientists and engineers. Furthermore, since students in the lab activity described here do not know the values of the physical properties they are measuring until after they turn in their lab reports, the laboratory activity described here helps diminish the confirmation bias and outright fudging that students often employ to enhance the success of their experiment and, therefore, earn a higher grade.

Table 1. Pulsars available on the author's YouTube channel, with periods omitted from view (Walkup, 2020a,b,c,d,e). The periods are rounded to the nearest 0.001 seconds given that ordinary timing devices available to students only measure to the 0.01 seconds. The sources for each pulsar sound file are also tabulated (Department of Astronomy Cornell University, 2006; Pulsar Central, 2000).

Label	Pulsar	Period(s)	Link	Source
Pulsar A	B2020+28	0.343	https://youtu.be/sfJDeKSa208	RadioSky.com
Pulsar B	B0329+54	0.716	https://youtu.be/DW108PjTRe4	RadioSky.com
Pulsar C	B0833-45	0.089	https://youtu.be/E7uA5Zn9R1Q	Cornell
Pulsar D	B0950+08	0.253	https://youtu.be/2D1RxKT3mxg	RadioSky.com
Pulsar E	B1933+16	0.359	https://youtu.be/HUA1tLCuWoM	Cornell

1.2 Pulsars

Pulsars are the remnants of massive stars that have collapsed under their own weight. The resulting supernovae leave behind neutron stars of immense density but much smaller size. Because the moment of inertia of a neutron star is dwarfed by that of the original star, conservation of angular momentum makes the neutron star spin at incredibly high rates. Spin frequencies range from 0.04 Hz for PSR J0250+5854 all the way up to 716 Hz for pulsar PSR J1748-2446ad (Hessels et al., 2006; Tan et al., 2018).

Pulsars emit beams of electromagnetic radiation along their magnetic axes. Because the axis of rotation of a pulsar does not necessarily correspond to its magnetic axis, this radiation functions similarly to a navigation beam for any observer that crosses its path. As such, the pulsar appears to "pulse" in much the same way that a lighthouse appears to pulse to someone on a ship.

Some astronomers have recently mapped the electromagnetic pulses of pulsars to sound and others have loaded these files onto YouTube for easy playing (Department of Astronomy Cornell University, 2006; Pulsar Central, 2000; s7Range, 2018) This allows students to "hear" the pulses of the pulsar with nothing more than an internet connection.

The goal of this lab activity is for students to identity a mystery pulsar by timing its period. Most sources of pulsar sound files also note the true periods of the pulsars, therefore negating their use as a mystery pulsar. In response, one of the authors (Walkup) has uploaded some of the clearer sound files to separate YouTube pages (Walkup, 2020b,c,d,e). These pulsars are listed in Table 1

2 Method

In this laboratory activity, students attempt to identify an unknown (at least to them) pulsar by timing its period, then entering their results into an online database. Although seemingly simple, their ability to obtain a result that exactly matches that of the unknown pulsar is effectively zero because of both random and systematic error. At best, they can only obtain a range of periods for which they can say the unknown pulsar "most likely" or "almost certainly" resides.

2.1 Reducing random error

The first step in this activity is to diminish random error by choosing a timing method that produces minimal variation in results. Using two different timing methods, students measure the standard deviation (Eqn.1) in each, where T_i represents the ith individual period measurement and $\bar{T}$ represents the average of all N measurements. Students simply choose that method which produces the smallest standard deviation.[1]

$$s = \sqrt{\frac{\sum_{i=1}^{N} (\bar{T} - T_i)^2}{N-1}} \qquad (1)$$

The laboratory assistant can offer any number of different timing methods to students. For example, students can compare the precision between using a stopwatch and the timer on their cell phones. Because we were forced to teach online, we could not expect students to possess a stopwatch, so we changed the methods to the following:

1. Measure the time between each pulse using 5 trials.
2. Measure the time it takes for 5 pulses to complete, then dividing by 5.

Students should note that the total time of measurement is the same in both cases; only the manner in which the time is divided differs between the two. In our lab, we relate this time of measurement in terms of observatory cost, that is, both methods will cost the astronomy team the same in terms of money and labour, but one would likely prove more precise than the other. In this sense, this first step offers a glimpse into data-driven decision making.

2.2 Reducing systematic error

In Step 2, students use the optimal method found in Step 1 to measure the period of a pulsar with a disclosed period. By comparing their sample mean to that of the known period, they determine the bias in their optimal method. This bias allows them to correct for systematic error through calibration. For example, if students obtained a period of 0.52 seconds for a pulsar of known duration 0.50 seconds, then they know to calibrate any future results by subtracting 0.02 seconds, in principle eliminating the systematic error in their measurements.

2.3 Estimating unknowns

In Step 3, students use the optimal method found in Step 1 to time the period of a mystery pulsar, correcting for the bias found in Step 2. They also calculate the standard error

$$SE = \frac{s}{N} \qquad (2)$$

in their results by hand or using a spreadsheet. They then translate this standard error into the language of confidence intervals straddling the sample mean of their results. For students new to statistics, we have found the following explanations helpful:

1 Strictly speaking, students should use the population standard deviation σ rather than the sample standard deviation s. However, since they are only comparing one standard deviation with the other, such a distinction serves no practical purpose.

1. The period of the unknown pulsar most likely ranges between T_{low} to T_{high}, where T_{low} and T_{high} represent 1 *SE* above and below the sample mean, respectively. Note that "most likely" corresponds to roughly a 68% likelihood.
2. The period of the unknown pulsar almost certainly ranges between Tlow to Thigh, where T_{low} and T_{high} represent 2 *SE* above and below the sample mean, respectively, where "almost certainly" corresponds to roughly a 95% likelihood.

Students then enter these ranges into the search tool provided by the ATNF Pulsar Catalogue using the variable P0, which represents the barycentric period of the pulsar (Manchester et al., 2005; Australia Telescope National Facility, 2020). The database then generates a list of pulsars that fall within these ranges. These pulsars serve as "likely culprits" for the mystery pulsar.

For example, students that obtain an average pulse of 0.044 seconds with a standard error of 0.001 seconds can assume the period of the mystery pulsar "most likely" falls between 0.043 and 0.045 seconds and "almost certainly" falls between 0.042 and 0.046 seconds[2]. Inserting the command "P0 > 0.043 && P0 < 0.045" into the text field titled "Condition" generates the three pulsar candidates in Table 2. Using a range generated by ±2 SE, we can generate the larger number of pulsars in Table 3 for which the unknown pulsar almost certainly is listed (assuming that students have effectively calibrated away systematic error in Step 2).

3 Results

A total of 33 students enrolled in a calculus-based introductory physics laboratory course completed the pulsar activity during the Northern Hemisphere Spring 2020 semester. The authors omitted 8 students from consideration because they displayed no meaningful effort to complete the activity, leaving 25 students in the sample. Of these, 5 were able to produce a list of pulsars containing the identity of the mystery pulsar.

Gauging success proved difficult. For example, a few students (5) failed to provide a list containing the mystery pulsar, but largely because their confidence intervals were so small that all but the tiniest bias pushed their range of periods too far from the true period. In this sense, the more careful they conducted Step 1, the less likely they would find suitable candidates for the mystery pulsar. Alternatively, sloppy timing would produce such large confidence intervals that a student's chances of producing a list containing the mystery pulsar would grow. As such, we graded the students' lab reports holistically based on their ability to generate reasonably small confidence intervals and reduce bias.

The number of pulsars students listed in the 68% confidence interval ranged from 11 to 17, with 21 to 30 in the 95% confidence interval [3]. Ten students relied on a mean period that was so far from the true that any meaningful reflection on their results would have uncovered an obvious mistake. with significant differences that could have been easily found (e.g.: it was easy to see that the measured period was way too large).

2 Out of concerns for brevity, we are using a relatively short pulse as an example in this article because there are relatively few pulsars that have such periods. For the same reason, we chose a standard error more precise than the students can obtain with a simple stop watch. As such, students will generate much a much larger list of pulsars than shown here.

3 One student only listed one pulsar, the correct one. We consider this result an anomaly.

Table 2. Pulsars generated by the ATNF Pulsar Catalogue with periods ranging ± 1 SE about the sample mean of 0.044 seconds, assuming SE = 0.001 seconds.

No.	Pulsar	Frequency (Hz)	Period (s)
1	J0557-2948	22.91337153767	0.0436426389
2	J1157-5112	22.94144832067	0.04358922706
3	J1813-1749	22.37171236	0.04469930526

Table 3. Same as Table 2, but using a confidence interval of ± 2 SE.

No.	Pulsar	Frequency (Hz)	Period (s)
1	J0453+1559	21.8427329517106	0.04578181687
2	J0557-2948	22.91337153767	0.0436426389
3	J1157-5112	22.94144832067	0.04358922706
4	J1813-1749	22.37171236	0.04469930526
5	J0453+1559	21.8427329517106	0.04578181687
6	J1454-5846	22.10004678013	0.045248773

4 Discussion

Although a challenge, results show that with enough care students can generate a reasonably small number of pulsars containing the identity of the mystery pulsar. This activity does not involve any more knowledge of astronomy than the lab instructor desires. As such, it poses a useful lab activity for the first few weeks of instruction and addresses many concepts in error analysis that students will need for the remainder of the semester. For example, this activity provides useful dialog in terms of significant figures, as students will try to input ridiculously precise values into the database. Typically, students fail to appreciate the importance of significant figures and treat them simply as a requirement they must satisfy so as to not lose scoring credit. In this activity, paying heed to significant figures changes the range of periods they submit to the ATNF database and therefore the number of pulsars the database will offer.

This activity also clarifies the difference between the standard deviation from the standard error. Students only use the standard deviation in Step 1 because at that point they only care which of the two methods produces results with the least variation. They rely on the standard error in Step 3 because they need an estimation of how far their sample mean will range from the true mean. This activity also clearly delineates random from systematic error because each type of error is examined in its own step of the process. Furthermore, the impact of random error and systematic clarify: Systematic error pushes the confidence interval away from the true sample mean. This means that their list might not include the mystery pulsar at all.

A large random error, however, generates a list too large. While the mystery pulsar may be listed in the results, the huge number of listed pulsars makes the results worthless. These considerations post challenges for any objective attempt to grade lab reports simply on whether students included a list with the mystery pulsar in it. We suggest holistically grading students' results based on two considerations: (a) Does the size of their confidence interval indicate careless timing in Step 1 and (b) does the bias in their results indicate substandard calibration in Step 2?

References

Australia Telescope National Facility (2020). The ATNF Pulsar Database. Retrived from https://www.atnf.csiro.au/research/pulsar/psrcat/.

Department of Astronomy Cornell University (2006). Pulsar sounds. Retrived from http://hosting.astro.cornell.edu/~deneva/psr_sounds/pulsars_sounds.htm.

Hessels, J. W., Ransom, S. M., Stairs, I. H., Freire, P. C., Kaspi, V. M., and Camilo, F. (2006). A radio pulsar spinning at 716 hz. *Science*, 311(5769):1901–904. https://doi.org/10.1126/science.1123430.

Manchester, R. N., Hobbs, G. B., Teoh, A., and Hobbs, M. A. (2005). The Australia Telescope National Facility Pulsar Catalogue. *The Astronomical Journal*, 129(4):1993. https://doi.org/10.1086/428488.

Pulsar Central (2000). Retrived from http://www.radiosky.com/rspplsr.html.

s7Range (2018). Pulsar sounds. Retrived from https://youtu.be/35SbvnYEc9c.

Tan, C. M., Bassa, C. G., Cooper, S., Dijkema, T. J., Esposito, P., Hessels, J. W. T., Kondratiev, V. I., Kramer, M., Michilli, D., Sanidas, S., and et al. (2018). Lofar discovery of a 23.5 s radio pulsar. *The Astrophysical Journal*, 866(1):54. https://doi.org/10.3847/1538-4357/aade88.

Walkup, J. R. (2020a). Pulsar A. Retrived from https://youtu.be/sfJDeKSa2O8.

Walkup, J. R. (2020b). Pulsar B. Retrived from https://youtu.be/DWl08PjTRe4.

Walkup, J. R. (2020c). Pulsar C. Retrived from https://youtu.be/E7uA5Zn9R1Q.

Walkup, J. R. (2020d). Pulsar D. Retrived from https://youtu.be/2D1RxKT3mxg.

Walkup, J. R. (2020e). Pulsar E. Retrived from https://youtu.be/HUAltLCuWoM.

Walkup, J. R., Key, R. A., and Talbot, P. R. (2019). Timing a light pulse in the introductory physics lab. *Physics Education*, 54(4):045007. https://doi.org/10.1088/1361-6552/ab15b4.

Astronomy Education Journal, 2021, Vol. 01, No. 1

doi: 10.32374/AEJ.2021.1.1.007

RESOURCES & ACTIVITIES

Resource guides for astronomy educators and their students

Andrew Fraknoi[1,*]

[1]Fromm Institute, University of San Francisco

[*]fraknoiandrew@fhda.edu

Abstract

This is a brief review of my long-term project to create annotated resource guides for astronomy educators and their students. The college-level guides focus on a range of astronomical topics, such as "Women in Astronomy," "Science Fiction and Astronomy," "Black Lives in Astronomy," "Music Inspired by Astronomy," "Light Pollution, Radio Interference, and Satellite Swarms," etc. Their aim is to bring together pointers to a range of reliable and interesting materials for instruction, discussion, or student projects in one convenient and easy-to-find place.

Keywords: Resources; Interdisciplinary; Bibliographies; Instructional Materials; Journals; Diversity

1 Origin and History of the Project

The resource guides began to take shape when the author was editor of Mercury magazine at the Astronomical Society of the Pacific (ASP) during the 1970's and 1980's. Rather than featuring individual book reviews, which other astronomical publications were already doing, we hit on the idea of reviewing introductory materials on a specific topic, and including magazine articles as well as books. This was before the World Wide Web existed, when finding relevant materials for teaching or outreach was more laborious.

Among these first efforts were:

- A Subject Index to Astronomy in Scientific American Magazine (Fraknoi, 1977b)
- Interdisciplinary Approaches to Astronomy (Fraknoi, 1977a)
- The Sun: A Reading List (Fraknoi, 1977c)
- Close Encounters with [Astronomical] Pseudoscience (Fraknoi, 1978)
- Computer Programs in Astronomy (Mosley and Fraknoi, 1985)

Eventually, the guides became independent of the magazine, and were printed (and, occasionally, specially written) for workshops for teachers (in the ASP's Universe in the Classroom series) and symposia for Astronomy 101 instructors (in the ASP's Cosmos in the Classroom series.) Later, some of those guides were collected in the ASP's The Universe at Your Fingertips, which became available as a DVD-ROM (`https://myasp.astrosociety.org/product/DV122/the-universe-at-your-fingertips-20-dvd-rom`).

The contents of the guides began to cover a wider range of subjects, including a longer resource covering the contributions of women to astronomy and a catalog of music inspired by astronomy.

In the decade of that began in 2011, some of the existing guides, and many new ones, were moved to the ASP's webpage (now at: `https://astrosociety.org/education-outreach/education-resources.html`) and also to the web page: `http://www.fraknoi.com/resource-guides-on-astronomy-education/`.

More recently, the most-often requested guides have been assigned convenient, short URLs, using the *bit.ly* notation. In the next section, we list and explain some of the guides that seem to be the most useful for educators. Bear in mind that all the guides focus on non-technical materials, so that they are accessible to both educators and students. And all the guides are limited to materials available in English.

2 Current Guides Available on the Web

2.1 Guides Related to Diversity

The Contributions of Women to Astronomy: http://bit.ly/astronomywomen gives resources on the general issues that women have faced (and continue to face) in astronomy, and then suggests books, articles, and webpage about the life and work of 37 women: 20 living scientists, and 17 noted women of the past. In addition, there are one or two resources for another 24 women about whom less non-technical material is available.

Black Lives in Astronomy: http://bit.ly/blackastro provides published and web resources on the lives and work of 26 black astronomers in North America, selected because material about them was easily available to students. It begins with a short section on the general topic of the challenges and triumphs of black astronomers.

The Astronomy of Many Cultures: http://bit.ly/astrocultures is designed to make U.S. students more familiar with the astronomical contributions and thinking of cultures outside North America and Europe. While it is no way complete, it does provide materials for an introduction to the astronomy of the indigenous people of Central and South America, Asia, Africa, Polynesia, and Australia.

2.2 Guides to Interdisciplinary Approaches

Science Fiction Stories with Good Astronomy: http://bit.ly/astroscifi is an annotated catalog of written science fiction that has reasonable science as its basis. It is organized by astronomy topic, so you can look up good novels and stories to recommend on black holes, Mars, or SETI, for example. A number of the short stories are now available on the Web and URLs are provided .

Music Inspired by Astronomy: http://bit.ly/astronomymusic is another annotated catalog, of pieces of music that are based on serious astronomy (and not just a casual mention of the Moon.) It includes both classical and popular music, and is organized by astronomy topic. It is limited to pieces that are available on disk or on the Web. Don't miss the Supernova Sonata, the Hubble Cantata, or the five rock and roll songs about black holes!

Plays and Films Inspired by Astronomy: http://bit.ly/astroplays is a shorter listing of selected plays and movies (in English) that feature astronomers as characters or revolve around astronomical ideas or history. This list is in its first editions and needs additional suggestions.

Responding to Astronomical Pseudo-science: http://bit.ly/astropseudoscience is a guide to resources to help astronomers or their students learn the skeptical, rational perspective about such "fiction sciences" as astrology, UFOs as alien spaceships, ancient astronauts, crop circles, creationism (or intelligent design), astronomical doomsdays, and much more.

(Note: In 2021, particularly in the U.S., an enormous amount of gullible media attention is being paid to military images and videos that are claimed to show mysterious spaceships with unexplainable behavior. But investigators have found earthbound explanations that are much more reasonable. A new one-page guide to web resources on these and other UFO issues is newly available at: http://bit.ly/ufoskeptic)

This Day in Astronomical History: http://bit.ly/astrodates is a month-by-month calendar of 158 astronomical anniversaries or events to enrich a class or talk, including a number involving work by women. (And, unlike most such calendars, the emphasis is not on human spaceflight anniversaries, but on real astronomical research.)

2.3 Guides for Teaching Astronomy

Doing Astronomy Outreach and Education When You Can't Go Out: http://bit.ly/astrooutreach is a recent compilation of resources for doing remote education and outreach (both by necessity during the pandemic, and by choice at other times.) The guide also includes other lists of resources created by veteran educators and by crowdsourcing.

Short Astronomy Videos to Go with Each Chapter of OpenStax Astronomy: http://bit.ly/shortastronomyvideos is an annotated listing of free videos in English, available on the web, each shorter than 10 minutes, to help illustrate and illuminate astronomy concepts. The catalog is keyed to the free, on-line, open-source, introductory astronomy textbook for which I am the lead author, but can be used with any other course materials as well. Many NASA and ESA videos are included and organized by topic.

Sources of Astronomy Images: http://bit.ly/astronomyimages is a brief, incomplete list of where beginners in astronomy education can find the largest catalogs of astronomy images to use in their work. I apologize that the emphasis is on sources where captions are available in English.

M.O.O.S.E: Menu of Outreach Opportunities for Science Education: http://bit.ly/AASMOOSE (all caps) is a web document, put together for the American Astronomical Society's Astronomy Ambassadors program, to provide early-career astronomers with resources, guidelines, and opportunities for doing outreach. The focus is on the US, but the majority of resources could be adapted for astronomy outreach in other countries.

Astronomy and Humor: http://bit.ly/astrohumoris a short, selective list of sources for astronomical jokes, cartoon, and other humorous materials – to spice up a talk, a final exam, or post-conference visit to a bar.

2.4 Miscellaneous Guides

Light Pollution and Dark Skies : http://bit.ly/darkskyguide is beginner's guide to issues where human activities are posing a problem for astronomical research. Designed, as all these guides are, so students can also use them, it includes nontechnical explanations of the problem and proposed solutions. Other topics included are radio interference and satellite swarms.

Pluto and the Kuiper Belt: http://bit.ly/plutokuiper is a 2019 guide to introductory readings about Pluto the dwarf planet, the controversy of its classification, and its exploration by New Horizons. It also focuses on the Kuiper Belt, the flyby of the KBO previously known as Ultima Thule (now officially called Arrokoth) and the first claims of a Planet Nine in the belt.

An Index to the ASP's Mercury Magazine: http://bit.ly/mercuryindex is a listing of all the main articles that appeared from 1972 through 2019 in the Astronomical Society of the Pacific's popular-level magazine. You can find collections of Mercury in all the libraries that have the journal *Publications of the ASP*. Many noted astronomers explained their work for the non-scientist in the pages of this magazine.

Recently, I have put together a one-page catalog with links to all my active resource guides in one convenient place. You can find it at: http://bit.ly/fraknoiguides

3 Other Guides and Future Efforts

Between 2001 and 2013, I served as Founding Co-editor, and later as editorial consultant, for Astronomy Education Review, a refereed journal on astronomy education, published first by the National Optical Astronomy Observatories and then by the American Astronomical Society. A subject index to all the published issues is available at: https://aas.org/teach/subject-index-papers-astronomy-education-review-2001-2013 It can serve as a good introduction to the state of astronomy education research in those years. A number of other guides, some now a few years old, are featured at the two web addresses given toward the end of section 1, above. I try to update each of the guides every few years or so, but it's a bit random when each one has its turn. I very much welcome suggestions for any of these guides. My hope is that they can, in some small way, make the efforts of astronomy educators more effective and point them in useful directions.

References

Fraknoi, A. (1977a). Interdisciplinary approaches to astronomy. *Mercury*, 6(5):20.

Fraknoi, A. (1977b). Subject index to astronomy in scientific american (1960-1976). *Mercury*, 6(1):20.

Fraknoi, A. (1977c). The sun: A reading list. *Mercury*, 6(4):20.

Fraknoi, A. (1978). Close encounters with pseudoscience. *Mercury*, 7(1):21.

Mosley, J. and Fraknoi, A. (1985). Computer programs in astronomy. *Mercury*, 14(1):27.

Astronomy Education Journal, 2021, Vol. 01, No. 1
doi: 10.32374/AEJ.2021.1.1.008

OPINION

Should we teach general relativity in high school? Why and how?

Jorge E. Horvath[1,*] and Pedro H. R. S. Moraes[1,2]

[1]Instituto de Astronomia, Geofísica e Ciências Atmosféricas, Universidade de São Paulo, Brazil ; [2]Universidade Federal do ABC (UFABC) - Centro de Ciências Naturais e Humanas (CCNH), Brazil

[*]foton@iag.usp.br

Abstract

We discuss the proposal of teaching General Relativity and its spin-offs (black holes, gravitational waves, cosmology) as a part of the High School curriculum, a task already undertaken by a few countries. We point out the importance of this proposal for the students and some possible reasons for the resistance of educators and scientists. A suggestion about how to circumvent the well-known mathematical and conceptual difficulties associated with this formidable task is made. It is not enough to stay at a qualitative level of explanation for High School students, although it would be desirable and positive at younger ages.

Keywords: Teaching General Relativity; Mathematical content; Secondary School

1 General Relativity, "Modern Physics" and Teaching

A strong argument for the 21st century as "the Century of Gravitation" can be made based on the advances in theoretical/observational Cosmology (Daniel et al., 2010; Ishak, 2019), study and imaging of real black holes (Akiyama et al., 2019; Eckart et al., 2017)(including the Nobel Prize that was recently awarded to three scientists for their work on the supermassive black hole at the Milky Way center (Schirber, 2020), discovery of gravitational waves (Abbott et al., 2016) and other important developments (Joshi and Malafarina, 2011; Kochanek et al., 2001; McKee and Ostriker, 2007; Metcalf et al., 2019). These wonderful facts, with frequent presence in the media, are unified by their fundamental Relativistic character, that is, Newtonian gravitation is not enough to support and explain the new findings. Our main point in this short Opinion text is that while it may be enough to introduce some qualitative material related to gravitation before High School, it is imperative to bring some quantitative/formal approach for High School students for a number of reasons including the very nature of Sciences. In the long run, we believe that a qualitative, conceptual approach leads to the loss of a fundamental connection with other branches of science. We elaborate on this point below.

It is no secret that under the time-dependent name of "Modern Physics", still widely used in schools and introductory university courses, groups a bunch of disciplines that are a century old or so. The lag between the educational content in hard sciences and the frontier knowledge shared by the scientists is enormous and has been put at the center of the debate many times (Goldader, 2002; Pasachoff, 2002). The rise of General Relativity (GR) as a concrete frontier topic adds a new twist to an old problem: *What should we teach? and how?*

Several colleagues would certainly agree with attempts to "close the gap". As a matter of fact, there are a handful of concrete initiatives to introduce these topics to catch up with the research news and give the students a broader perspective of science and the knowledge of the world. South Korea, Norway, Scotland (Kersting et al., 2018) are nationally engaged, and other countries like Australia feature Centres of Excellence and people dedicated to follow this path (Blair et al., 2016). We concretely target High School students, mature enough to be attained by quantitative and conceptual arguments, some of them possibly on their way to a career in science. In this sense, the material developed for younger students is very valuable since each age group needs a different instructional approach, and a qualitative one would be excellent for them.

Closing this gap is definitely important, but not only for "catching up" purposes. For a High School student to have a real glimpse of some GR concepts, as well as their applicability, could be of extraordinary and unparalleled value for his/her genuine interest in science, even for those that will be heading to different activities.

Here, we are not talking about a highly non-trivial spacetime metric and, say, "non-ideal" energy-momentum tensor, to compute the Einstein tensor components, derive the Einstein field equations and even solve them for all variables involved. We are, for example, encouraging the teaching of the metric (tensor) first notions, perhaps by quoting that the metric tensor is what rules the measuring of distances in curved spaces. The introduction of Minkowski metric is very welcome as an elegant shortcut to Euclidian space.

We also believe a deep approach of Einstein's equivalence principle is wholesome for High School students, bringing to students the possibility of wondering about distinguishing between "purely acceleration" and gravitation effects is certainly of great value. This could also prompt them to search for a profound understanding of coordinate systems and coordinate transformations. Finally, the tidal forces calculation should inform the learner if they are dealing with a false (due simply to coordinate transformations) or genuine gravitational field.

A closer examination to the material and topics presented to the students about GR gravitation reveals that the vast majority is of qualitative character, and also that new approaches to the education (like the use of cartoons, videos, etc.) are widely used for this task. It seems that there is a consensus on the difficulties of concepts of GR topics, and an implicit recognition that a formal approach is out of question. In other words, that "the mathematics is not there till we put it there" (A.S. Eddington) and that concepts precede the mathematical description. We face here a first problem: many "hard" scientists would feel that the essential part of gravitation is lost without any mathematical formulation. In fact, this can be traced back to an old resonance among scientists of Galileo's words:

> *"The universe] cannot be read until we have learnt the language and become familiar with the characters in which it is written. It is written in mathematical language, and the letters are triangles, circles and other geometrical figures, without which means it is humanly impossible to comprehend a single word." (G. Galilei, Il Saggiatore, p.175)*

and the very evolution of hard sciences as well, which can be qualified as "hermetical", that is, need preparation on specific tools and ideas to advance.

In fact, the relation of mathematics and the rest of "hard sciences" seems to be based to a good extent on a kind of symbiosis, in which the formal mathematical formulation is not optional, but rather is a part of the *logos* (McDonnell, 2017). These kinds of beliefs are in fact the basis of modern science curricula for hard sciences. While this has never been a problem, for example, for biological sciences, the "softening" of physics and astronomy, even with the purpose of making accessible difficult subjects, would prompt some degree of rejection among the practitioners. As an example, it is clear that the survival of basic topics in High School curricula is based on the conviction that they are important, a doorway to more complicated things, but also on the fact that they are "doable" mathematically, even if restricted to simple exercises and problems, for example, kinematics, thermal physics and other well-known topics.

We believe that if teaching GR implies giving up mathematics as part of the *logos*, this would prompt a long-term change in the very concept of hard sciences, something seen as undesirable to say the least. It is not too early to pretend an exposure of High School students to a set of carefully designed quantitative topics related to GR.

2 Which is the path forward?

Having gauged the situation of gravitational physics, we must address the question of *what to do with education*?. There are a number of resources not yet explored for the construction of a viable approach to GR and other difficult subjects, in which not only the concepts but also the formal formulations are difficult, such as quantum mechanics. Recently, the application of Bruner's ideas (Bruner, 2003)to a narrative contribution to GR has been suggested (Cardoso and Gurgel, 2013), as identified in Einstein's methodology. *Gedankenexperiments* are quite related to this task. In addition, the exploration of a less formal approach to the mathematical content present, for example in the famous work by Misner, Thorne and Wheeler (2017)2017. We believe this latter approach, when properly adapted to the High School level, could be a bridge between purely conceptual teaching and a kind of "math without math", pointed to recover the very essence of the symbiosis between physics and mathematics.

The whole idea is a kind of "verbal mathematics" first, aimed to illuminate the deep meaning of relations between physical quantities. An example is Newton's Law written as:

$$\textit{variation of velocity} \propto \textit{force} \tag{1}$$

followed by a proper discussion and formalization leading to $\vec{F} = m \times \vec{a}$ that is, giving time and tools to grasp a more abstract form of physical reasoning through mathematics. Such a path can be also followed towards GR expressions, such as the renowned field equations written as:

$$\textit{curvature of spacetime} \propto \textit{distribution of matter/energy} \tag{2}$$

at least those involving scalar and vector quantities. Of course, in this case it is mandatory for the students to previously perceive the space-time concept and its non-absolute character, possibly through the well-known rubber sheet experiment (one can also get in touch with Reference (Stachel, 2005)), stressing to the students the difficulty in curving space with a simple numerical estimate. Overall, a kind of "computational thinking" would be developed by students and the main ideas will be encoded by some symbols, ultimately, this is what mathematics is all about.

The connection with "Newtonian analogues" could also be explored. In fact, a "quasi-Newtonian" approach would be an important step forward for a variety of important topics, such as black holes and gravitational waves. For example, Friedmann equations were written in a Newtonian mood (Misner et al., 2017, p. 708). The quasi-Newtonian derivation of the Schwarzschild radius $R = \frac{2GM}{c^2}$ (Visser, 2005) is another example. Finally, we note that well-suited material, involving pre-calculus and geometry targeted to these tasks exists (Kraus and Zahn, 2018; Schutz, 2003; Zahn and Kraus, 2014, 2018) and would be an excellent starting point to conduct large-scale educational research aimed to test how it leads to significant learning (Ausubel, 2000) and what can be done to improve its efficiency in the classroom.

In short, it is clear that a long-term investigation of this possibility and its testing with professors and students is ahead and, in many cases, already available, a feasible yet involved pedagogical strategic project with ample applicability. The important point to be stressed again is not to give up a minimal degree of formalization without trying.

As a final statement, we strongly agree with the introduction of General Relativity (and other topics) in High School, but believe that a qualitative, conceptual approach only causes a loss of a fundamental connection with other branches of science. It is sometimes observed that a procedure like this is enough to "separate", for example, astronomy from physics, as if fundamentally different things. Is it enough to teach astronomy and physics qualitatively than nothing at all? Yes, at least in elementary school, but we should also attempt to develop a viable approach suited to this new challenge for the High School level. This is possible and would be important to retain the hard science character of these disciplines, while also reinforcing the unity of sciences in a concrete fashion. A hard science taught "without mathematics" is not an achievement we should be proud of.

References

Abbott, B. P., Abbott, R., Abbott, T., Abernathy, M., Acernese, F., Ackley, K., Adams, C., Adams, T., Addesso, P., Adhikari, R., et al. (2016). Observation of gravitational waves from a binary black hole merger. *Physical review letters*, 116(6):061102. http://dx.doi.org/10.1103/PhysRevLett.116.061102.

Akiyama, K., Alberdi, A., Alef, W., Asada, K., Azulay, R., Baczko, A. K., Ball, D., Baloković, M., Barrett, J., et al. (2019). First m87 event horizon telescope results. i. the shadow of the supermassive black hole. *Astrophysical Journal Letters*, 875(1):L1. https://doi.org/10.3847/2041-8213/ab0ec7.

Ausubel, D. P. (2000). *The Acquisition and Retention of Knowledge: A Cognitive View*. Netherlands: Kluwer Academic Publishers.

Blair, D., Henriksen, E., and Hendry, M. (2016). Why don't we teach einstein's theories in schools. *The conversation*. https://theconversation.com/why-dont-we-teach-einsteins-theories-in-school-69991.

Bruner, J. (2003). *Making Stories: Law, Literature, Life*. Netherlands: Kluwer Academic Publishers.

Cardoso, D. and Gurgel, I. (2013). A complementaridade das linguagens narrativas e matemática no contexto da elaboração da relatividade geral. In *XX Simpósio Nacional de Ensino de Física – SNEF 2013 – São Paulo, SP (in portuguese)*, São Paulo, Brazil.

Daniel, S. F., Linder, E. V., Smith, T. L., Caldwell, R. R., Cooray, A., Leauthaud, A., and Lombriser, L. (2010). Testing general relativity with current cosmological data. *Physical Review D*, 81(12):123508. http://doi.org/10.1103/PhysRevD.81.123508.

Eckart, A., Hüttemann, A., Kiefer, C., Britzen, S., Zajaček, M., Lämmerzahl, C., Stöckler, M., Valencia-S, M., Karas, V., and García-Marín, M. (2017). The milky way's supermassive black hole: How good a case is it? *Foundations of Physics*, 47(5):553–624. http://doi.org/10.1007/s10701-017-0079-2.

Goldader, J. (2002). Letter to the editor: Comment on jay pasachoff's article,"what should college students learn?". *Astronomy Education Review*, 1(2). https://doi.org/10.3847/AER2002012.

Ishak, M. (2019). Testing general relativity in cosmology. *Living Reviews in Relativity*, 22(1):1–204. http://doi.org/10.1007/s41114-018-0017-4.

Joshi, P. S. and Malafarina, D. (2011). Recent developments in gravitational collapse and spacetime singularities. *International Journal of Modern Physics D*, 20(14):2641–2729. http://doi.org/10.1142/S0218271811020792.

Kersting, M., Henriksen, E. K., Bøe, M. V., and Angell, C. (2018). General relativity in upper secondary school: design and evaluation of an online learning environment using the model of educational reconstruction. *Physical Review Physics Education Research*, 14(1):010130. http://doi.org/10.1103/PhysRevPhysEducRes.14.010130.

Kochanek, C., Keeton, C., and McLeod, B. (2001). The importance of einstein rings. *The Astrophysical Journal*, 547(1):50. http://doi.org/10.1086/318350.

Kraus, U. and Zahn, C. (2018). Sector models—a toolkit for teaching general relativity: Iii. spacetime geodesics. *European Journal of Physics*, 40(1):015602. https://doi.org/10.1088/1361-6404/aae3b5.

McDonnell, J. (2017). The role of mathematics in fundamental physics. In *The Pythagorean World*, pages 69–124. Springer. https://doi.org/10.1007/978-3-319-40976-4_3.

McKee, C. F. and Ostriker, E. C. (2007). Theory of star formation. *Annu. Rev. Astron. Astrophys.*, 45:565–687. https://doi.org/10.1146/annurev.astro.45.051806.110602.

Metcalf, R. B., Meneghetti, M., Avestruz, C., Bellagamba, F., Bom, C. R., Bertin, E., Cabanac, R., Courbin, F., Davies, A., Decencière, E., et al. (2019). The strong gravitational lens finding challenge. *Astronomy & Astrophysics*, 625:A119. https://doi.org/10.1051/0004-6361/201832797.

Misner, C. W., Thorne, K. S., and Wheeler, J. A. (2017). *Gravitation*. NJ, USA: Princeton University Press.

Pasachoff, J. M. (2002). What should college students learn. *Astronomy Education Review*, 1(1):124–130. https://doi.org/10.3847/AER2001012.

Schirber, M. (2020). Nobel prize: Facing the reality of black holes. *Physics*, 13:158. https://physics.aps.org/articles/v13/158.

Schutz, B. (2003). *Gravity from the Ground Up: An Introductory Guide to Gravity and General Relativity.* Cambridge UK: Cambridge University Press.

Stachel, J. (2005). Development of the concepts of space, time and space-time from newton to einstein. In Ashtekar, A. A., editor, *100 years of Relativity*, pages 3–20. Singapore: World Scientific.

Visser, M. (2005). Heuristic approach to the schwarzschild geometry. *International Journal of Modern Physics D*, 14(12):2051–2067. https://doi.org/10.1142/S0218271805007929.

Zahn, C. and Kraus, U. (2014). Sector models—a toolkit for teaching general relativity: I. curved spaces and spacetimes. *European Journal of Physics*, 35(5):055020. https://doi.org/10.1088/0143-0807/35/5/055020.

Zahn, C. and Kraus, U. (2018). Sector models—a toolkit for teaching general relativity: Ii. geodesics. *European Journal of Physics*, 40(1):015601. https://doi.org/10.1088/1361-6404/aae3b7.