Cross-sectional study of students' knowledge of sizes and distances of astronomical objects

Vinesh M. Rajpaul,[1,2] Christine Lindstrøm,[3,4,*] Megan C. Engel,[2]
Morten Brendehaug,[3] and Saalih Allie[5]

[1]*University of Cambridge, Astrophysics Group, Cavendish Laboratory,*
J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom
[2]*Department of Physics, University of Oxford, Oxford OX1 3RH, United Kingdom*
[3]*Faculty of Education and International Studies, Oslo Metropolitan University,*
PB 4 St. Olavs plass, N-0130 Oslo, Norway
[4]*Centre for Computing in Science Education, Department of Physics, University of Oslo,*
Sem Sælands vei 24, N-0316 Oslo, Norway
[5]*Department of Physics and Academic Development Programme, University of Cape Town,*
Rondebosch, South Africa, 7701

This paper reports on the results from administering a modified version of the Introductory Astronomy Questionnaire (IAQ) to middle school students and preservice science teachers in Norway. Ranking tasks formed a key part of the instrument, and we detail a new method for analyzing ranking task data. One of our main findings was that even after instruction, a significant proportion of students held erroneous views and conceptions regarding sizes, distances, and the nature of basic astronomical entities, such as stars and planets. We argue that the commonalities between some of the issues we identified and those presented in extant studies—from a variety of countries, with samples ranging from junior high school students and undergraduates to primary school teachers—may point to deeper cognitive issues inherent in, and possibly unique to, engaging with astronomy.

I. INTRODUCTION

A challenge all individuals face in acquiring basic knowledge of astronomical objects is that we have skewed, limited, or no direct experience with these objects. In mechanics—which has been the focal point of considerable research in physics education research (PER) [1,2]—we observe phenomena in a direct, experiential manner; then, via ontological innovation [3], we describe the essential core of the phenomenon. On the other hand, the directly observable aspects of astronomy are only very weak proxies for the inferred phenomena: the Earth is the largest object with which a human has any direct experience (anyone who has been on an intercontinental flight can attest to the enormous size of Earth); the $\sim 100\,000$ km/h orbital motion of Earth around the Sun is all but imperceptible; the Sun is a disc of modest proportions in our field of view; and observing stars as twinkling points of light in

the night sky does little to suggest the gargantuan sizes, high velocities, and extremely energetic processes associated with these stars—or the vast distances between them [4]. Indeed, the actual scales of astronomical phenomena and objects lie, almost without exception, far beyond human experience, and the little experience we do have—with Earth, the Sun, the Moon, and stars—provides a skewed starting point for building basic knowledge of these and other astronomical objects.

In this study, we explore students' basic knowledge of astronomical objects on scales ranging from telluric (i.e., pertaining to Earth as a planet) to cosmological, focusing primarily on their perspectives of relative sizes and distances. In so doing, we address the intersection of two under-researched topics [5] in astronomy education research (AER): students' knowledge of *size and distance* and of *modern topics* concerning astronomy beyond the Solar System [6–8].

II. BACKGROUND

Previous studies of size and distance have all been relatively modest parts of larger studies that have primarily focused on the most familiar astronomical objects—Earth, Sun and Moon—across many different samples (students

---

*christine.lindstrom@hioa.no

and educators at all levels of education). Because size and distance have not been at the forefront of these studies, little effort has been made to enable comparison with other studies. The following three studies illustrate the point. Bakas and Mikropoulos [9] investigated 102 Greek students 11–13 years old and found, using multiple choice questions (MCQs), that 81% of the students knew that the Sun is larger than Earth. Cin [10] found through interviews with 65 Turkish middle school students (aged 14) that 43% had a correct perception of the relative sizes of the Sun, Earth, and Moon, but it is not possible to extract how many knew the relative sizes of only the Sun and Earth. Summers and Mant [11] asked 120 British primary school teachers to select the most representative MCQ option of relative size of Earth compared to the Sun in a scale model. 87% of the teachers chose one of the five alternatives in which Earth was smaller than the Sun (32% chose the correct option), only 2% chose the equal or larger option, and 12% responded "don't know." Beyond the general finding that a significant number of students at all levels do not have correct notions of relative sizes of the most well-known objects in the Solar System, it is difficult to compare results across studies because the results are strongly linked to how the questions were asked.

Even within the same sample, the way a concept is probed strongly affects the result. In a MCQ survey given to 1414 high school students from around the United States, Sadler [12] included two different questions probing student understanding of the distance between Earth and the Moon: one question showed five drawings of the two objects and asked students to select the scale model that best reflected Earth and the Moon; the other question asked students to select from five values the distance in miles that most closely represented the distance between the two objects. Whereas the correct numerical distance was the most popular option (selected by 30% of students), only 13% selected the correct scale model; 40% chose a drawing in which the Moon was only 2–3 Earth diameters away. This mismatch between results from the same construct probed in different ways was also found by Cin [10] in interviews with Turkish students: although many students ranked the Sun, Earth, and Moon correctly in terms of size, when they were then prompted to *draw* a picture comprising these three objects, the relative sizes did not in general reflect their previous statements.

The lack of standardization in *how* to probe students' understanding of sizes and distances makes it difficult to compare—and sometimes even interpret—results. MCQs are often unavoidably leading, such as Summers and Mant's question, in which many more alternatives where Earth was *smaller* than the Sun were offered. Others are not entirely clear, such as Bakas and Mikropoulos' phrasing of "Which celestial body is bigger, the Earth or the Sun, and how many times bigger?" where the only two

options for relative sizes they provided were 10 or 1 000 000 times bigger, without mentioning to which measure (radius, volume, etc.) this size referred.

Whereas a multiple choice format can be suitable for comparing relative sizes of two or three objects, it becomes impractical with four or more objects as the researchers are forced to reduce the ranking alternatives to a handful of options, thus restricting the students in expressing their actual beliefs [13]. Interviews do not have the same restrictions, but while they enable the students to freely discuss their thoughts on relative size and distance of multiple objects, it is a time-consuming methodology that severely limits the number of participants, does not enable identical questioning, and is subject to interviewer expertise and bias [14].

III. RESEARCH FOCUS

A. Research question 1

Previous studies have produced specific results in terms of percentage of students holding certain notions relating to size and distance, many focusing on detailed knowledge of exact distance or scale relationship. However, in terms of pedagogical usefulness, we argue that as a first step it is more important to know *whether* students know that one object is larger than another before delving into details of *how much* larger the object is. Given the wide variety of question styles across studies in the literature, it was also difficult to compare the relative knowledge of different student samples. Hence, our first research question (RQ) was the following.

RQ1: What are the most prevalent incorrect views among students regarding sizes and distances of astronomical objects, and how do these differ—if at all—across levels of education?

B. Research question 2

In the literature, there is an expressed assumption that an understanding of size and distance (e.g., in scale models) in astronomy is needed to be able to explain astronomical phenomena [5,15]. However, there seem to be no empirical results supporting this intuitive expectation—in fact, Fanetti [15] failed to find a correlation between students' knowledge of the scale of the Earth-Moon system and ability to explain lunar phases. Instead, we hypothesized that basic knowledge of astronomical objects—comprising both qualitative *and* ranking knowledge—might be a better predictor of students' ability to explain a wide variety of astronomical phenomena than knowledge of size and distance alone (regardless of whether it is measured as ranking or scale knowledge). One reason for this hypothesis was that whereas students may be able to rank a pair of objects correctly by chance—thus introducing noise into the data—the extent to which they can supply good qualitative explanations for objects should be far less

susceptible to random guessing. To test this hypothesis, we formulated our second RQ, which could inform future work involving size and distance.

RQ2: Is there a correlation between students' qualitative and ranking knowledge of astronomical objects?

IV. METHODOLOGY

A. Context and sample

Our study samples were (i) students drawn from eight middle schools in Norway, and (ii) preservice teachers at the largest teacher education institution in Norway, in both cases before and after astronomy instruction.

Norway is a European country with a population of just over 5 million. It has one of the world's highest per capita incomes, and often ranks first in the world on indices measuring human development, well-being, prosperity, and democracy. Its welfare model includes universal health care, a comprehensive social security system, and public education, which is effectively free at all levels, regardless of nationality [16–19].

The Norwegian education system comprises elementary school (years 1–7), middle school (years 8–10), and high school (years 11–13). More than 97% of elementary and middle school students are enrolled in public schools [20], where year 10 marks the end of compulsory education. Science is compulsory in years 1–11, whereas in years 12 and 13, physics, chemistry, and biology—which are separate subjects in these years—are noncompulsory. Science contains one astronomy module in middle school, generally covered in year 8.

There is significant latitude for individual year 8 teachers to choose astronomy topics on which to focus and test; for example, national guidelines simply note that students should be able to "describe the universe and different theories for how it has evolved" and to "investigate a topic from the exploration of space, synthesize and present information from different sources." Nevertheless, at the year 8 level, almost all teachers elect to use one of just three physics textbooks [21–23], leading to a level of curriculum uniformity across schools.

The year 12 physics course features astronomy in its modern physics curriculum, which covers stellar astronomy (Stefan-Boltzmann law, Wien's displacement law, HR diagrams, and the life cycle of stars) and the standard model for the evolution of the Universe. Given that year 12 students write a national examination, the physics curriculum is standardized across schools.

Norwegian post-secondary education is offered in universities and university colleges (the distinction between which today is largely historical). The higher education system is in accordance with the Bologna process, with three-year bachelor's degrees and two-year master's degrees, although some degree programs deviate from this norm by taking four years to complete.

Because of the dominance of public school education in Norway, preservice science teachers represent a self-selected subset of the public school educated populace. To qualify as a middle school science teacher in Norway, the most common route before 2017 was to complete a four-year Bachelor of Teaching degree, which included the equivalent of one year of science (Science I and II). Middle school teachers are also qualified to teach in upper elementary school (years 5–7). Preservice elementary school science teachers are required to take Science I, whereas Science II is optional. Consequently, the cohort enrolled in Science II comprises a mixture of future middle and elementary school teachers. As of 2017, however, teacher education has moved to a compulsory five-year model. All preservice teachers who began their studies in 2017 or later are now required to complete a joint bachelor's *and* master's degree, with the latter including a research project.

The 41 (24 females and 17 males) preservice teachers in this study all attended the largest teacher education institution in Norway, viz. Oslo and Akershus University College (HiOA). They had passed Science I and were enrolled in Science II in the 2013–2014 academic year. The students belonged to two separate cohorts: one cohort was in their third year and did Physics II in Fall 2013, whereas the other cohort was in their fourth year and did Physics II in spring 2014. C. L. was the physics instructor for both cohorts. Both courses comprised the subjects physics, chemistry, biology, technology & design, meteorology & geology, and science education. Although these subjects are nationally determined, the detailed curriculum of each subject varies between institutions. Physics I and II at HiOA constitute approximately 20% of Science I and II. Three consecutive 45-min lessons are delivered in each 3-h class. Physics I comprises seven such classes (plus one class used for group presentations) and covers thermodynamics, gravity & buoyancy, sound, light, kinematics, forces and energy; Physics II comprises ten classes and covers electromagnetism (electricity, magnetism, and induction), atomic and nuclear physics (atomic physics, nuclear physics, and radiation physics) and astronomy (the Sun-Moon-Earth system, the Solar System, and the universe, plus one outdoor observing night). The HiOA guidelines for the science teacher education program are quite brief in their description of the content of the science courses; the content of the astronomy module in Physics II is described as "the Solar System, the universe, the evolution of the universe." It is up to each individual physics lecturer to devise a list of learning goals.

As for the middle school students, a total of eight nonrandomly selected schools in and around Oslo (the Norwegian capital) agreed to participate in the study during fall 2014. None of the 535 year 8 students and all 387 of the year 10 students in these schools had studied the semistandardized astronomy module, so the year 8 students

were considered the preinstruction sample, and the year 10 students (from the same schools), the postinstruction sample. Of the 904 (out of 922) middle school students that elected to specify their gender, 452 were male and 452 were female.

Note that whereas the pre-post comparisons for the preservice teachers represented a within-group comparison, for the middle school sample it was a between-group comparison. Given the large numbers of middle school students participating in our study, we aggregated all year 8 students to form the preinstruction sample, and all year 10 students to form the postinstruction sample: regardless of any fine-grained differences in the composition of these two samples, the essential distinguishing feature was that no students in the former sample would have received any astronomy instruction, whereas all students in the latter sample would have studied the year 8 astronomy module. The samples were otherwise homogeneous in terms of student background, gender composition, and schools attended; any instructor effects and individual student variability should have been averaged out by the large sizes of the samples.

Our rationale for studying both middle school students and pre-service teachers was the following. As noted earlier, middle school marks the highest level at which all Norwegian students receive compulsory astronomy instruction, while after receiving tertiary astronomy instruction, the preservice teachers will go on to teach science, including astronomy, to future generations of Norwegian school students. Thus, we expected that any misconceptions identified among secondary-level students might be better understood by investigating whether their instructors shared any of these difficulties.

B. Ranking tasks

Given the lack of suitable approaches in the literature, we first sought a method for conveniently probing a large sample of students' conceptions of size and distance of several (5–10) astronomical objects. A simple ranking task appeared an ideal choice: it can easily be included as a survey item, requires a minimum of writing, is unambiguous in its interpretation, and is easy to digitize for analysis. As a written test item (as opposed to an interview item), it ensures that all respondents receive identical questions, and written surveys are cheap and easy to administer [14].

Ranking tasks are powerful for probing sizes and distances because they test neither memorized answers nor students' abilities to manipulate formulas blindly, and allow multiple simultaneous comparisons from one question [24]. From previous studies, we can conclude that a majority of students (but far from all) know the relative sizes of the Sun and Earth. However, a ranking task does not get bogged down with details such as how many times bigger or further one object is compared to another, which

is only relevant *after* students have established *which* is bigger or further; hence our emphasis on *"ranking knowledge"* in this study (a more comprehensive knowledge of size and distance would include knowledge of magnitude, which we would term *"scale knowledge"*). By comparing many objects, ranking tasks allow us to go beyond a simple statement of what percentage of students know the correct ranking of a particular pair, by enabling us to identify what the most prevalent incorrect ideas are among a range of pairwise comparisons. The latter holds much greater power as it provides information relevant for directing pedagogical efforts.

Ranking tasks have been used in physics since the 1980s [25], but they have almost exclusively been used in problem sets, not as test items. A 5-item ranking task admits a total of $5! = 120$ possible different responses, while a 10-item task allows a total of 3.6×10^6 possible different responses, with each task having only one correct solution. (If students are constrained to a MCQ format for ranking tasks, they may typically be limited to only four or five possible responses; cf., e.g., Slater *et al.* [26]) Thus, despite their apparent simplicity, such ranking tasks can lead to a very broad spectrum of responses, from which a wealth of potentially interesting data may be extracted.

From the literature, it was not clear what a reasonable upper limit for the number of items in a ranking task was, so we explored this from a theoretical and an experimental perspective. Theoretically, it may be shown that there exist procedures for sorting (ranking) items that require at worst $\log_2 n!$ operations, and in best cases only n operations [27]. An expert who has developed an appropriate mental model might reasonably approach such efficiency by (unconsciously) using efficient, comparison-based sorting heuristics, and could thus sort a list of 10 items using perhaps 10 or 20 mental comparisons. For example, they might readily recognize that the edge of the observable universe must be the most distant item on a given list, or that items residing outside of the Solar System must be more distant than those within it, without actively having to compare these to every other item. Even in the worst case, however, comparing every single item with every other item in the list requires $n(n-1)/2$ (in the case of a 10-item list, 45) comparisons. For $n = 10$, this is not an unreasonable number, given that the full list could be skimmed in a matter of seconds; for values of n much larger than 10, however, inefficient or inexpert approaches to sorting would certainly become unfeasible. Experimentally, we piloted a 10-item ranking task (introduced later in the paper) with about a hundred undergraduate students enrolled in an introductory astronomy course at the University of Cape Town (UCT). When we discussed the ranking task with this pilot group, it was clear that students interpreted the task as intended, and that incorrect responses usually stemmed from incorrect conceptions of the relative positions of the astronomical objects in question, rather than from random guessing [28].

C. The NIAQ instrument

The Norwegian Introductory Astronomy Questionnaire (NIAQ) is an instrument developed to evaluate certain aspects of the astronomy knowledge of Norwegian preservice science teachers, with elements suitable for middle school students. It was adapted and translated from the original Introductory Astronomy questionnaire (IAQ), developed by Rajpaul *et al.* [29] for an introductory astronomy course; details of the adaptation and motivation for doing so are described in Lindstrøm *et al.* [30].

The focus of this paper is on three of the questions in the NIAQ: (Q5a) size rankings and (Q5b) simple explanations of five different astronomical objects, viz. galaxy, planet, star, universe, solar system; and (Q7) ranking in terms of distance from Earth's surface of ten different items, chosen to span many orders of magnitude, including the Sun, the center of the Milky Way, the edge of the observable Universe. The full questions constituting the NIAQ, in both English and Norwegian, appear in Appendix A and B. Whereas Q5a and Q5b were part of the original IAQ, the 10-item ranking task (Q7) was *not*. This new ranking task was developed subsequent to the publication of the original IAQ study, based specifically on experience with difficulties encountered by students taking the introductory astronomy course at UCT [30], so that the NIAQ enabled exploration of students' conceptions of size *and* distance on a cosmological scale. (Two forthcoming papers in this series will focus on the analysis and results of the remaining questions in the NIAQ.)

D. Administering the NIAQ

The entire NIAQ instrument was given to the preservice science teachers, whereas a subset of four questions was given to the middle school students. The questions given to the middle school students were chosen based on their relevance to the expected knowledge of middle school students, as well as their relative ease of marking.

The NIAQ was administered to both cohorts of preservice teachers as a pretest during the atomic and nuclear physics module and as a post-test after the main astronomy module (this was after the Science II examination for the fall cohort, but before the examination for the spring cohort). Students were informed of the purpose of the questionnaire, were told that it would not affect their course marks in any way, and were given 45 min to complete it. All preservice teachers who were present when the NIAQ was administered elected to complete the questionnaire; the pre- and post-tests were completed by 40 and 38 preservice teachers, respectively. No differences were found between the cohorts during the analysis, so they are not separated in the results.

For the middle school administration, year 8 students were given 25 min and year 10 students 20 min to complete the questionnaire. All questionnaires were administered by M. B., who informed the students that the questionnaire was part of his Master's project and that completing the questionnaire would be of great help to him but would not affect the students' marks in any way. All students who were given the questionnaire returned it.

The response rate was 97%–100% for all questions, for both the middle school students and the preservice teachers.

E. Analyzing the NIAQ

1. Ranking tasks (Q5a and Q7)

There were no suitable suggestions in the literature for how to optimize the extraction of relevant information from ranking tasks. While it is straightforward and common practice to assign to each student an overall score for a ranking task [31] or for an item in a ranking task, such simple descriptive statistics would not paint a complete picture. Far more information could be teased out by analyzing detailed relationships between ranks assigned to different objects in a given task. Consequently, the ranking tasks (Q5a and Q7) were analyzed using a novel method developed specifically for this data set.

For the purposes of illustration, suppose we had asked students to rank the numbers "three, five, one, two, four" from biggest to smallest. First, each student's full response was captured on a computer (e.g., 1 for one, 2 for two, etc.); thus, a correct response might be captured as $\{1, 2, 3, 4, 5\}$, while a response that omitted the item in rank three might be captured as $\{1, 2, 4, 5, -\}$. Once all responses had been captured, we wrote computer code to accomplish the following tasks.

First, we checked all responses for validity. A response was regarded as invalid if it (a) assigned a single rank to more than one item, e.g., $\{1 \,\&\, 3, 2, 4, 5, -\}$, or (b) assigned multiple ranks to a single item, e.g., $\{1, 2, 1, 4, 5\}$, or (c) omitted any of the required items, e.g., $\{-, 2, -, -, 1\}$. Thus we avoided having to devise complex and perhaps arbitrary schemes for evaluating valid responses on an equal footing with responses that were either ambiguous, as in cases (a) or (b), or incomplete, as in case (c). Subsequent analysis was restricted to valid responses. We then assigned a score to each response: i.e., the number of items ranked correctly. In order tease out extra information from the full spectrum of responses, however, we supplemented this basic scoring with two further, more detailed analyses.

First, we calculated how many responses placed a given item in each possible position in the ranking, which enabled a more nuanced evaluation of the incorrect responses. For example, $\{*, *, 2, *, *\}$ might be considered a less serious mistake than $\{*, *, *, *, 2\}$, even though the item "two" is ranked incorrectly in both cases (here we use the asterisk as a wildcard character, to denote any of the other items in the response).

Second, for each valid response, we also computed the relative ranks of all possible *pairs* of items, and recorded all mistakes (e.g., ranking $1 > 2$, ranking $2 > 4$, etc.). This yielded further valuable information about students'

attempts to rank items, even in cases where absolute rankings were wrong. For example, $\{5, 1, 2, 3, 4\}$ has no items ranked correctly, yet correctly captures the relations $1 < 2 < 3 < 4$, albeit while wrongly implying that $5 < 1$. On the other hand, $\{5, 4, 3, 2, 1\}$ does not have any items ranked correctly even in a relative sense.

2. Explaining task (Q5b)

Question 5b prompted students to provide brief explanations, at a level suitable for their peers, of the following astronomical objects: galaxy, planet, star, universe, solar system (lack of capitalization deliberate). Note that these were the same objects students were asked to sort in the ranking task immediately preceding this explaining task.

A subset of the responses was translated independently by C. L. and by another native Norwegian speaker who is fluent in English; V. R. cross-checked the translations for consistency, and any small ambiguities or disagreements were discussed and resolved. Subsequently, the remaining responses were translated by C. L. only. More details about the translation are provided in Lindstrøm *et al.* [30].

The marking scheme used for the translated responses was exactly the same as that used for the identical question in the original IAQ [29]; in short, we sought to probe whether students had a qualitatively correct understanding of the entity in question, and one which they could communicate to someone else, rather than whether they could produce a detailed technical explanation. Therefore responses were marked as incorrect (0 points), partially correct (0.5 points), or minimally correct or adequate (1 point), and blank responses were not assigned scores. Criteria for an explanation to qualify as minimally correct (1 points) are given below.

- **Galaxy:** a collection or system of stars and other material, and any information to distinguish it from, e.g., a stellar system or star cluster (e.g., student mentions "billions of stars").
- **Planet:** an object in orbit around the sun (or another star), and any piece of information to distinguish it from, e.g., an asteroid or comet (larger than a certain size, stable due to its own gravity, cleared its immediate neighborhood, could have its own moons orbiting it, etc.).
- **Star:** a large or massive, hot or luminous sphere of plasma or ball of gas, or any equivalent explanation.
- **Universe:** all existing matter and space, all of the cosmos, everything, the totality of existence, a connected space-time, or any equivalent explanation.
- **solar system:** the Sun and the objects in orbit around it (e.g., planets, moons), *or* that it is a system comprising one or a small number of stars that orbit each other.

Explanations that only partially matched the above criteria for a given object were awarded 0.5 points (e.g., "a planet is any body which orbits a star"), as were students who provided examples without further explanation (e.g., "a star is something like the sun"). Responses that did not match the above criteria and/or were factually incorrect were awarded 0 points (e.g., "a planet is any place that can support life").

The preservice teachers' explanations were marked independently by V. R. and C. L.; agreement in the scores assigned was found to be $> 95\%$, with Cohen's kappa coefficient of $\kappa \approx 0.92$. Given the very high interrater reliability achieved with the smaller sample of preservice teachers, the middle school students' explanations were marked by C. L. only.

3. Correlation analyses

We used "off-the-shelf" statistical software to study correlations between scores on the ranking and explaining tasks, e.g., by computing linear or rank correlation coefficients and associated p values; conducting chi-squared tests of independence between scores from the two tasks; etc. In our case, we used MATLAB's Statistics and Machine Learning Toolbox to carry out these analyses, though we note that equivalent open-source packages are freely available online [32].

V. RESULTS

This section is organized such that we address RQ1 first, by reporting on the results of each of the three questions in the NIAQ; afterwards we present a correlation analysis to address RQ2.

A. Ranking task: sizes of astronomical objects (Q5a)

The results from this question, which asked students to rank five objects—galaxy, planet, star, universe, solar system—from smallest to largest, are summarized in Figs. 1, 2, and 3. The three most noteworthy results are the following.

(1) *Preservice science teachers were more knowledgeable than middle school students, though neither sample exhibited gains in knowledge following instruction* (Fig. 1). The average score (number of items ranked correctly out of total number of items) obtained on this task by the middle school students was 72.5% preinstruction and 69.6% postinstruction; this decrease is not statistically significant ($p = 0.15$, using a two-sample t test). The average score for the preservice teachers was 90.5% preinstruction, and 91.6% postinstruction; this gain is not statistically significant ($p = 0.82$). However, the preservice teachers' scores were significantly higher than the middle school students' scores ($p \ll 0.001$). We also note that both pre- and postinstruction, more than 70% of students in all samples could correctly rank the Solar System, a galaxy, and the universe (Fig. 2).
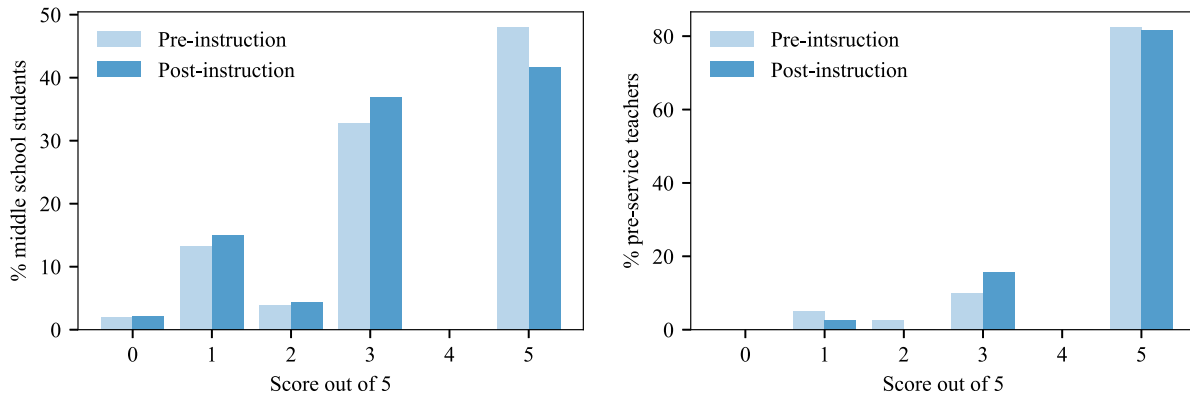
FIG. 1. Distribution of scores (number of items ranked correctly) for the ranking task in Q5a, for middle school students (left panel) and preservice teachers (right panel). The average score for the middle school students was 72.5% preinstruction and 69.6% postinstruction; the average score for the preservice teachers was 90.5% preinstruction and 91.6% postinstruction.

(2) *The dominant incorrect view among middle school students was that planets are bigger than stars* (Fig. 3). Both pre- and postinstruction, more than 40% of middle school students did not seem to know that a planet was the smallest item on the list, and indeed more than 40% of middle school students specifically ranked a planet as being bigger than a star [33]. Fewer than about 10% of preservice teachers made the latter mistake, though this was nevertheless the most common mistake postinstruction.

(3) *The second most commonly held incorrect view among middle school students was that galaxies are smaller than solar systems* (Fig. 3). Nontrivial numbers of middle school students (15%–20%) responded that the Solar System is larger than a galaxy. Similarly, 15% of preservice teachers before instruction seemed to think the same, but this incorrect idea was largely rectified after instruction (only 5% in the post-test).

B. Explaining astronomical objects (Q5b)

The results from this question are tabulated in Tables I and II for the middle school students and the preservice teachers, respectively. The most notable results are the following.

(1) *The Universe and solar system were the best known items to all samples of students.* All middle school students (both pre- and post-instruction) fared best on "universe" (average score ∼70%) and "solar system" (∼55%), with average scores of around 30% for the other objects. As with the middle school students, the preservice teachers fared best on universe and solar system.

(2) *Significant numbers of students are aware of the existence of exoplanets, i.e., planets around stars other than the Sun.* Though not directly related to the research questions, a notable fraction of middle school students clearly interpreted solar system (capitalization deliberately avoided in question

(a)	(1)	(2)	(3)	(4)	(5)
Planet	59	37	2	1	1
Star	40	56	4	1	0
Solar system	1	4	81	13	1
Galaxy	0	4	12	76	8
Universe	1	0	1	8	90

(b)	(1)	(2)	(3)	(4)	(5)
Planet	58	37	4	1	1
Star	40	54	5	1	0
Solar system	1	5	74	18	2
Galaxy	1	4	15	73	7
Universe	1	0	2	7	90

(c)	(1)	(2)	(3)	(4)	(5)
Planet	90	8	0	3	0
Star	10	90	0	0	0
Solar system	0	0	88	13	0
Galaxy	0	3	13	85	0
Universe	0	0	0	0	100

(d)	(1)	(2)	(3)	(4)	(5)
Planet	92	8	0	0	0
Star	8	87	5	0	0
Solar system	0	5	89	5	0
Galaxy	0	0	5	92	3
Universe	0	0	0	3	97

FIG. 2. Size ranks assigned to astronomical objects in Q5a, by middle school students (a) preinstruction and (b) postinstruction and preservice teachers (c) preinstruction and (d) postinstruction. Each matrix element indicates the percentage of respondents who assigned a specific rank (column number) to a specific item (row). The objects are ordered so that the diagonal corresponds to correct rankings, i.e., planet should be ranked first, star second, etc. The color scale is added simply to aid visual clarity, with the level of green saturation highlighting items most often ranked correctly, and the level of red saturation highlighting the most prevalent mistakes. For example, matrix (a) tells us only 59% of the preinstruction middle school students could identify a planet as the smallest item in the list, regardless of how they ranked other items, while 13% of them wrongly thought that a solar system (or the Solar System) was the second largest item in the list. The numbers of *valid* responses analyzed to produce each matrix, out of the total number of nonblank responses, were (a) 519 out of 522, i.e., 99%, (b) 372 out of 384, i.e., 97%, (c) 40 out of 40, i.e., 100%, and (d) 38 out of 38, i.e., 100%, respectively.

(a)	Planet	Star	Solar system	Galaxy	Universe
Planet	-	41	3	3	1
Star	-	-	4	2	1
Solar system	-	-	-	15	2
Galaxy	-	-	-	-	9
Universe	-	-	-	-	-

(b)	Planet	Star	Solar system	Galaxy	Universe
Planet	-	42	3	4	1
Star	-	-	5	4	1
Solar system	-	-	-	20	4
Galaxy	-	-	-	-	9
Universe	-	-	-	-	-

(c)	Planet	Star	Solar system	Galaxy	Universe
Planet	-	10	3	3	0
Star	-	-	0	0	0
Solar system	-	-	-	15	0
Galaxy	-	-	-	-	0
Universe	-	-	-	-	-

(d)	Planet	Star	Solar system	Galaxy	Universe
Planet	-	8	0	0	0
Star	-	-	5	0	0
Solar system	-	-	-	5	0
Galaxy	-	-	-	-	3
Universe	-	-	-	-	-

FIG. 3. Percentages of students who made specific pairwise errors on the size ranking task in Q5a; the sample groups (a)–(d) match those in Fig. 2. The elements of the matrix should be interpreted as follows: value in row i and column j = percentage of respondents who wrongly ranked item i as being larger than item j. The lower triangular part of the matrix is empty, since all pairwise relations here correspond to correct relative rankings. For example, matrix (a) tells us that 41% of the preinstruction middle school students wrongly ranked a planet as being larger than a star, regardless of the absolute ranks they assigned to either object or to any other objects, while 9% of them thought that a galaxy is larger than the universe. A color scale is superimposed on the elements of the matrix simply to highlight the most prevalent mistakes.

phrasing) as "stellar system" in their explanations (21% preinstruction and 19% postinstruction), without making reference to the Sun itself or the planets in our Solar System. Notably, a *majority* of preservice teachers (75% preinstruction, and 61% postinstruction) interpreted solar system as stellar system in their explanations, and many preservice teachers explicitly interpreted "planet" to mean "exoplanet."

(3) *Middle school students showed no change pre- to postinstruction, whereas preservice teachers started from a higher baseline and demonstrated a clear*

increase in knowledge of all five items. Among the middle school students, *no* statistically significant differences were found between the average pre- and postinstruction scores for any of the five objects. Taking all five objects into account, the average score both pre- and postinstruction was 40%. Among the preservice teachers, by contrast, the mean scores for all objects increased pre- to postinstruction. The biggest gain was seen on "star" (19 percentage points), and the smallest gain (5 percentage points) on universe. Indeed, when analyzed on a

TABLE I. Middle school students' scores on the explaining task (Q5b), for both the pre- and postinstruction samples. For each object, the total number of students answering the question, the number of students scoring 0, 0.5, or 1 out of 1, and the mean score for the object (arithmetic mean over all students' scores for that object) is shown.

	Preinstruction ($n = 518$)					Postinstruction ($n = 377$)				
Object	n(total)	n(1)	n(0.5)	n(0)	Mean score	n(total)	n(1)	n(0.5)	n(0)	Mean score
Planet	497	26	255	216	31%	344	11	182	151	30%
Star	489	52	202	235	31%	357	52	146	159	35%
Solar system	487	197	133	157	54%	348	149	83	116	55%
Galaxy	434	65	141	228	31%	300	40	98	162	30%
Universe	473	262	127	84	69%	341	206	74	61	71%

TABLE II. Preservice teachers' scores on the explaining task (Q5b), both pre- and postinstruction. For each object, the total number of students answering the question, the number of students scoring 0, 0.5, or 1 out of 1, and the mean score for the object is shown.

	Preinstruction ($n = 40$)					Postinstruction ($n = 38$)				
Object	n(total)	n(1)	n(0.5)	n(0)	Mean score	n(total)	n(1)	n(0.5)	n(0)	Mean score
Planet	40	17	15	8	61%	37	15	20	2	68%
Star	40	18	14	8	63%	37	27	7	3	82%
Solar system	40	30	4	6	80%	38	34	2	2	92%
Galaxy	40	12	17	11	51%	38	14	21	3	64%
Universe	40	31	5	4	84%	38	32	4	2	89%

TABLE III.   Correct solution to the ranking task in Q7b, along with approximate distances to the items in question (the question did not ask students to give the distances to the objects; these are included here for reference only, and to illustrate the many orders of magnitude in distance spanned by the items). Note: 1 AU (astronomical unit) equals approximately 1.5×10^8 km; 1 ly (light year) equals approximately 63 000 AU. The distance to the end of the Solar System depends on the definition used, but is not more than a few ly.

Rank	Item	Approximate distance
1	Ozone layer	20–30 km
2	Center of Earth	~6370 km
3	The Moon	~384400 km
4	The Sun	1 AU
5	Asteroid belt	2–3 AU
6	Neptune	29–31 AU
7	End of Solar System	50 AU to >1 ly
8	Pole star	433 ly
9	Center of Milky Way	~26000 ly
10	Edge of observable universe	~93 billion ly

student-by-student basis, almost every individual student's scores on the explaining task increased pre- to postinstruction. Taking all five objects into account, the average preservice teacher score increased from 69% to 78%.

C. Ranking task: distances to astronomical objects (Q7)

This question asked students to rank ten objects—center of the Milky Way; edge of the observable Universe; the asteroid belt; edge of the Solar System; the Moon; the Sun; the Pole star; the ozone layer; center of Earth; and Neptune—in terms of their *distance from Earth's surface*. When devising this question, the items were chosen to cover many orders of magnitude in distance. The choices were also motivated by experience with students' incorrect ideas about the size of Earth and how far away "space" is

(hence the inclusion of the ozone layer and the center of Earth), and with conceptualizing astronomical bodies in three dimensions. For reference, the correct ranking of the 10 items, along with approximate distances from Earth's surface, is given in Table III.

The results from this question are summarized in Figs. 4, 5, and 6. There is a wealth of data contained in these figures, and we outline below just a few of the most noteworthy results.

(1) *Pre-service science teachers were generally more knowledgeable than middle school students, and instruction only made a difference with the preservice teachers* (Figs. 4 and 5). The average score (number of items ranked correctly out of total number of items) for the preservice teachers increased from 52.8% preinstruction to 64.8% postinstruction, with this increase approaching statistical significance ($p = 0.056$). The average score obtained on this task by the middle school students was 35.3% preinstruction and 37.3% postinstruction, which is not a statistically significant difference ($p = 0.19$, using a two-sample t test). The preservice teachers' scores were significantly higher than the middle school students' scores ($p \ll 0.001$).

(2) *One of the most significant and persistent incorrect views was ranking the ozone layer as further away from the surface of Earth than the center of Earth* (Fig. 6). Both pre- and postinstruction, more than 50% of middle school students thought that the ozone layer is further away from Earth's surface than is the center of Earth. Fewer than a third—but still an alarmingly high number—of preservice teachers made the same mistake (31% preinstruction and 21% postinstruction). Taken at face value, this mistake suggests that many students thought the height of Earth's atmosphere is greater than Earth's radius. This conclusion is supported by conversations carried out by V. R. when piloting the question at UCT, prior to
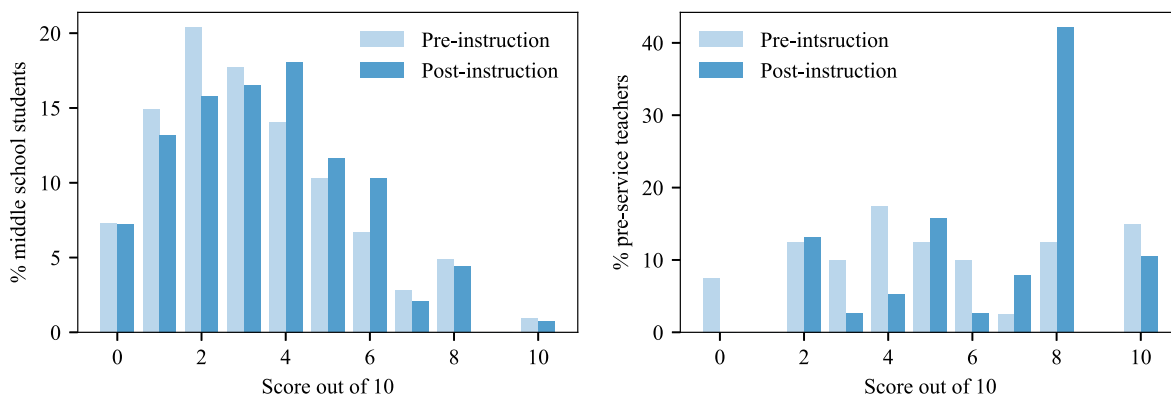

FIG. 4.   Distribution of scores (number of items ranked correctly) for the ranking task in Q7, for middle school students (left panel) and preservice teachers (right panel). The average score for the middle school students was 35.3% preinstruction and 37.3% postinstruction; the average score for the preservice teachers was 52.8% preinstruction, and 64.8% postinstruction.

(a)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ozone layer	36	40	4	3	3	4	3	3	3	0
Center of Earth	53	31	6	3	2	2	1	1	0	0
The Moon	5	14	58	11	5	3	2	0	1	0
The Sun	0	5	7	31	25	12	9	5	2	1
Asteroid belt	2	2	4	12	15	16	21	17	11	2
Neptune	1	2	7	13	25	26	12	9	4	1
End of Solar System	0	0	1	2	2	14	23	21	33	4
Pole star	1	2	9	15	12	12	14	22	14	1
Center of Milky Way	1	2	3	9	11	11	15	19	25	3
End of Universe	0	0	1	0	0	0	1	3	7	87

(b)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ozone layer	42	44	3	1	1	2	2	1	2	2
Center of Earth	51	38	4	2	2	1	2	0	1	0
The Moon	3	7	68	12	5	3	0	1	1	0
The Sun	1	3	4	27	31	17	8	5	2	1
Asteroid belt	1	2	3	14	14	13	17	18	16	2
Neptune	1	2	5	19	23	27	11	8	3	2
End of Solar System	0	0	1	2	3	14	25	20	31	4
Pole star	0	2	8	15	11	11	17	21	14	1
Center of Milky Way	1	2	4	6	9	10	17	22	25	4
End of Universe	0	0	0	0	2	1	2	4	6	85

(c)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ozone layer	69	26	0	3	0	0	0	0	3	0
Center of Earth	28	62	10	0	0	0	0	0	0	0
The Moon	3	8	79	8	3	0	0	0	0	0
The Sun	0	0	5	41	33	10	0	5	5	0
Asteroid belt	0	0	3	21	18	13	15	15	15	0
Neptune	0	0	3	28	23	36	8	3	0	0
End of Solar System	0	0	0	0	0	23	49	10	13	5
Pole star	0	5	0	0	18	10	18	38	10	0
Center of Milky Way	0	0	0	0	5	8	10	26	46	5
End of Universe	0	0	0	0	0	0	0	3	8	90

(d)	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Ozone layer	79	18	3	0	0	0	0	0	0	0
Center of Earth	18	76	5	0	0	0	0	0	0	0
The Moon	0	5	89	3	0	3	0	0	0	0
The Sun	0	0	0	26	50	21	3	0	0	0
Asteroid belt	0	0	0	58	21	8	3	8	3	0
Neptune	0	0	0	5	13	61	11	8	3	0
End of Solar System	0	0	0	0	3	3	68	11	16	0
Pole star	3	0	3	3	5	3	11	61	13	0
Center of Milky Way	0	0	0	5	8	3	5	13	66	0
End of Universe	0	0	0	0	0	0	0	0	0	100

FIG. 5. Distance ranks assigned in Q7 to astronomical objects by middle school students (a) preinstruction and (b) postinstruction and preservice teachers (c) preinstruction and (d) postinstruction. The matrices here should be interpreted in the same way as those in Fig. 2. For example, matrix (a) tells us that only 36% of the preinstruction middle school students knew that the ozone layer is closest item to Earth's surface, while a third of them wrongly thought that the end of the Solar System was the second-most distant item from Earth's surface. The numbers of valid responses analyzed to produce each matrix, out of the total number of nonblank responses, were: (a) 440 out of 512, i.e., 86% (b) 313 out of 376, i.e., 83% (c) 39 out of 39, i.e., 100%, and (d) 38 out of 38, i.e., 100%, respectively.

the present study: when shown a photograph of Earth from space, where it was clear that only a very thin blue line of atmosphere was visible above the enormous sphere of Earth itself, most students who answered the question incorrectly immediately expressed surprise or awe, and realized their mistake.

(3) *Another significant and persistent incorrect view was the belief that the Pole star resides* within *our Solar System* (Fig. 6). Both before and after instruction, more than 60% of middle school students and over a quarter of the preservice teachers seemed to think that the Pole star is closer to Earth than the edge of the Solar System—and one in three middle school students placed the Pole star closer to Earth than the Sun. This implies that many students

thought the Pole star is contained *within* the Solar System. More than one in five middle school students also placed the center of the Milky Way galaxy closer to Earth than the Sun.

(4) *Students exhibited confusion about the order of objects between the Moon and the end of the Universe.* More than half of students across all samples could assign correct ranks to the Moon and to the edge of the observable Universe (Fig. 5). These were, by significant margins, the two items that students were most often able to rank correctly. The objects between these two items, however, appear to be largely unknown to the middle school students [as witnessed by the considerable red shading of these objects in Figs. 6(a) and 6(b)],

(a)	Ozone layer	Center of Earth	The Moon	The Sun	Asteroid belt	Neptune	End of Solar System	Pole star	Center of Milky Way	End of Universe
Ozone layer	-	60	21	16	12	15	9	15	11	3
Center of Earth	-	-	13	7	6	6	2	6	6	2
The Moon	-	-	-	7	10	10	3	14	9	2
The Sun	-	-	-	-	32	29	8	33	24	3
Asteroid belt	-	-	-	-	-	62	31	49	42	4
Neptune	-	-	-	-	-	-	8	37	32	3
End of Solar System	-	-	-	-	-	-	-	64	57	6
Pole star	-	-	-	-	-	-	-	-	38	4
Center of Milky Way	-	-	-	-	-	-	-	-	-	5
End of Universe	-	-	-	-	-	-	-	-	-	-

(b)	Ozone layer	Center of Earth	The Moon	The Sun	Asteroid belt	Neptune	End of Solar System	Pole star	Center of Milky Way	End of Universe
Ozone layer	-	56	12	10	6	8	7	9	9	3
Center of Earth	-	-	10	6	4	5	3	4	4	2
The Moon	-	-	-	8	9	10	3	14	8	2
The Sun	-	-	-	-	33	39	7	33	22	3
Asteroid belt	-	-	-	-	-	63	32	55	44	4
Neptune	-	-	-	-	-	-	9	36	28	5
End of Solar System	-	-	-	-	-	-	-	60	53	8
Pole star	-	-	-	-	-	-	-	-	36	5
Center of Milky Way	-	-	-	-	-	-	-	-	-	7
End of Universe	-	-	-	-	-	-	-	-	-	-

(c)	Ozone layer	Center of Earth	The Moon	The Sun	Asteroid belt	Neptune	End of Solar System	Pole star	Center of Milky Way	End of Universe
Ozone layer	-	31	5	3	3	3	3	5	3	0
Center of Earth	-	-	10	0	0	0	0	3	0	0
The Moon	-	-	-	8	3	3	0	3	0	0
The Sun	-	-	-	-	31	44	5	15	10	0
Asteroid belt	-	-	-	-	-	56	33	44	28	0
Neptune	-	-	-	-	-	-	3	21	8	0
End of Solar System	-	-	-	-	-	-	-	38	26	8
Pole star	-	-	-	-	-	-	-	-	15	0
Center of Milky Way	-	-	-	-	-	-	-	-	-	5
End of Universe	-	-	-	-	-	-	-	-	-	-

(d)	Ozone layer	Center of Earth	The Moon	The Sun	Asteroid belt	Neptune	End of Solar System	Pole star	Center of Milky Way	End of Universe
Ozone layer	-	21	0	0	0	0	0	3	0	0
Center of Earth	-	-	5	0	0	0	0	3	0	0
The Moon	-	-	-	0	0	0	3	5	3	0
The Sun	-	-	-	-	61	16	3	11	11	0
Asteroid belt	-	-	-	-	-	13	8	16	13	0
Neptune	-	-	-	-	-	-	3	18	18	0
End of Solar System	-	-	-	-	-	-	-	26	24	0
Pole star	-	-	-	-	-	-	-	-	21	0
Center of Milky Way	-	-	-	-	-	-	-	-	-	0
End of Universe	-	-	-	-	-	-	-	-	-	-

FIG. 6. Percentages of students that made specific pairwise errors on the distance ranking task in Q7; the sample groups (a)–(d) match those in Fig. 5, while the matrices here should be interpreted in the same way as those in Fig. 3. For example, matrix (b) tells us that 60% of the preinstruction middle school students wrongly ranked the end of the Solar System as being further away from Earth's surface than the Pole star, regardless of the absolute ranks they assigned to either object or to any other objects, while a third of them wrongly thought that the Sun is further away from Earth's surface than is the Pole star.

and a source of much confusion—albeit significantly reduced after instruction—for the preservice science teachers [Figs. 6(c) and 6(d)].

(5) *A nontrivial number of middle school students thought that stars and objects in our Solar System are closer to Earth's surface than the ozone layer* (Fig. 6). The majority of preservice teachers knew that both the ozone layer and the center of Earth are much closer to Earth's surface than the other objects. However, before instruction, nontrivial and arguably alarming numbers of the middle-school students seemed to think that the Moon (21%), the Sun (16%), Neptune (15%), and even the Pole star (15%) are all closer to Earth's surface than the ozone layer. After instruction, these decreased to around 10% in each case. Taken at face value, these results suggest

that some students thought that the aforementioned objects might reside *within* Earth's atmosphere.

We also note that fewer than a quarter of students across all samples could correctly rank the asteroid belt (Fig. 5). More than 60% of middle school students and more than half of the preservice teachers before instruction seemed to think that the asteroid belt lies beyond Neptune; this number decreased to 13% for the preservice teachers, after instruction (Fig. 6). However, given the relatively similar distances between Earth and the Sun (1 AU) and between Earth and the asteroid belt (2–3 AU), we do not interpret incorrect ranking of the asteroid belt as a particularly grievous error; indeed, perhaps only "experts" might reasonably be expected to rank the item correctly.

D. Correlation between ranking (Q5a, Q7) and explanation of astronomical objects (Q5b)

We conducted a series of Pearson's chi-squared (χ^2) tests of statistical independence between scores on the two ranking tasks (Q5a, Q7)—as measured by the number of objects they ranked correctly—and their overall scores for the explaining task (Q5b). For the middle school students, we concluded that the association between scores on Q5a and Q5b was highly statistically significant: $\chi^2(2, N = 922) = 144.56$, with $p \ll 0.001$. Similarly, we found the scores on Q7 and Q5b to be very strongly associated: $\chi^2(2, N = 922) = 299.61$, and $p \ll 0.001$. For the preservice teachers, we found again a highly significant association between the scores for Q5a and Q5b: $\chi^2(2, N = 78) = 74.24$, with $p \ll 0.001$. This particular test did not, however, allow us to draw any decisive conclusions about an association between their scores for Q7 and Q5b, as we computed $\chi^2(2, N = 78) = 84.81$, with $p = 0.14$.

In fact, we found strong and statistically significant *linear* correlations between all scores on the two ranking tasks and their overall scores for the explaining task. These correlations existed both pre- and postinstruction for both the middle school students and the pre-service teachers, and in every case we found a linear correlation coefficient of $\rho > 0.4$, with $p \ll 0.001$. In some cases, the strength of the relationship increased when we measured it in terms of rank (e.g., using a Spearman or Kendall rank correlation coefficient) rather than linear correlation—presumably because rank correlation coefficients are less sensitive to the discrete "bunching" of scores around zero or full marks [34].

To probe further the way in which ability to rank objects may be correlated with descriptive knowledge of the objects, we homed in on the two object pairs associated with the most mistakes: star vs planet, and solar system vs galaxy. As our aim was simply to identify whether the correlation held for these specific pairs of items, regardless of how it might evolve over time, we aggregated responses from the year 8 and year 10 students for the middle school

analysis; similarly, we combined pre- and post-test teacher responses. Responses in each sample were then grouped into three categories, listed below.

- **No knowledge of at least one object.** Students who scored *incorrect* on at least one object in the ranked pair, regardless of the score for the other object. This category was based on the assumption that students with no knowledge of one of the objects in a pair lacked the knowledge basis from which to make a comparison, regardless of knowledge of the other object.
- **Partial knowledge of both objects.** Students who showed a minimum of partial knowledge of both objects, but did not show good understanding of both objects.
- **Good understanding of both objects.** Students who received full marks for their explanations of both objects were expected to be in the strongest position to correctly rank the pair of objects.

Table IV shows the results for the middle school students and the preservice teachers, respectively. The results confirm the assumption and provide details of performance.

(1) *The more knowledgeable students were about the nature of specific astronomical objects, the better they were at ranking them correctly.* Middle school students who displayed no knowledge of either planets or stars performed no better than chance in their ranking of these objects. Performance improved with increasing knowledge of the objects, a trend that generally also held for the pre-service teacher responses.

(2) *Students who could correctly define both objects in a pair had a near perfect record of correctly ranking the objects.* Of the 46 middle school students who correctly defined both objects in a pair, only two ranked the pair incorrectly (96% correct). The same trend was seen among the preservice teachers, with *none* of the 52 teacher responses in Category 3 ranking the objects incorrectly (100% correct).

We also note that the results are quite complex when comparing pairs of objects and student groups studied: for

TABLE IV. Correlating students' responses to Q5a and Q5b. Responses to Q5a were placed into three categories according to the quality of explanations provided for the objects in a ranked pair; for each category, the fraction of students who correctly ranked the objects (planet smaller than star, i.e., P < S; solar system smaller than galaxy, i.e., SS < G) is displayed.

	Middle school				Preservice			
	n	P < S	n	SS < G	n	P < S	n	SS < G
No knowledge	321	52%	267	81%	17	88%	18	78%
Partial knowledge	184	75%	149	97%	31	84%	37	89%
Good knowledge	10	90%	36	97%	29	100%	23	100%

the middle school students there are notable differences between the two pairs of objects, which is not the case for the preservice teachers; and while the teachers perform notably better than the middle school students for each category for planet vs star, this is not the case for solar system vs galaxy.

VI. DISCUSSION

The aim of this study was to explore students' basic knowledge of astronomical objects on scales ranging from telluric to cosmological, focusing primarily on their understanding of relative sizes and distances. To do so, we administered the NIAQ, or part thereof, to (i) preservice teachers at the largest teacher education institution in Norway, and (ii) students drawn from eight middle schools in Oslo. The preservice teachers were given the entire NIAQ, while the much larger sample of middle-school students were given an easy-to-mark subset of questions from the full NIAQ instrument.

Before receiving any postsecondary astronomy instruction, the preservice teachers represented a (self-selected) group of students who had received secondary-level astronomy instruction within the past few years; while after receiving their postsecondary astronomy instruction, the same preservice teachers would go on to teach science, including astronomy, to future generations of Norwegian school students. Our study thus contained both longitudinal and cross-sectional components.

As noted in Sec. IV, we could find no suitable suggestions in the literature for how best to extract from the ranking tasks all the information we were interested in. Specifically, the information we wished to extract was (i) the distribution of students' absolute ranks assigned to individual objects, and (ii) students' conceptions about the sizes and distances of each object to be ranked *relative to every other object.*

We presented this information using two different matrix representations. The first matrix presents the percentage of students who placed each object in a particular *position;* here one can see, at a glance, how well students fare on ranking specific objects, and whether students think a given object is smaller or larger (or closer or further away) than it really is. The second matrix presents the pairwise comparison of all possible pairs of objects, only showing the percentage of students ranking the order of the objects incorrectly (since the percentage of correct results is the complement). This enables comparison between any two objects, irrespective of a students' conceptions of any of the other objects or the absolute ranks assigned to the objects in the pair. The second matrix also allows comparison with results from other studies that include fewer or different objects. In tandem, these two matrices provide far more information than any scalar summary scores.

The main challenge for those who would like to replicate such an analysis is that there are no straightforward software packages that will produce such matrices. Therefore we hope to make open-source software [35] available that can be used for analysis of ranking tasks with arbitrary subjects (e.g., microscopic size scales, energy scales, process sequences such as stellar evolution or evolution of the early Universe, etc.) and different numbers of objects.

In hindsight, the inclusion of the asteroid belt in our longer ranking task may not have been the most prudent choice, given the similar distances between Earth and the Sun (1 AU) and between Earth and the asteroid belt (2–3 AU), and given that we did not attempt to link knowledge of the asteroid belt's location to knowledge of the formation of the Solar System, or of the distinction between the rocky and gaseous planets. Perhaps the "Andromeda Galaxy" would be a more sensible object to include in future iterations of the question, though this would depend on what one wished to study. On the other hand, it was unproblematic to include a challenging object in the ranking task: the preservice teachers showed a significant improvement in their ability to rank this object, and the poor performance by the middle school students was irrelevant when performing pairwise comparisons between any two objects other than the asteroid belt. For the same reason, if future iterations of the question were *not* to include the asteroid belt, straightforward comparisons with the present results would still be possible, further illustrating a virtue of our analysis method.

A. Prevalent incorrect views in students' ability to rank astronomical objects in terms of size and distance

Inspecting the large number of results presented in the matrices, we found that *some of the most significant and persistent incorrect views related to the most familiar celestial objects: Earth and stars.* Specifically, we found that three of the most dominant and persistent misconceptions were that (i) planets are bigger than stars, (ii) the ozone layer is more distant from the surface of Earth than is the center of Earth, and (iii) another star viz. the Pole star lies *within* our Solar System. The patterns of errors among the preservice science teachers mirrored those of the middle school students, although the preservice teachers were generally more knowledgeable about all objects, especially after instruction. We elaborate on these notable misconceptions below.

When comparing the relative sizes of planets and stars, only 60% of our middle school sample ranked planets as smaller than stars, whereas 90% of the preservice teachers knew the correct relative size. In the South African IAQ study, the majority of the undergraduate student sample knew that planets are larger than stars, both pre- and postinstruction (86% and 97%, respectively)—a result comparable to the Norwegian tertiary sample. No other studies have explored this particular generalized question, so the most relevant comparisons are with studies that

probed students' conceptions of the relative sizes of Earth and the Sun. Our middle school students performed worse than 11–13 year olds studied by Bakas and Mikropoulos [9], 81% of whom knew the correct relative size of Earth and the Sun; however, this may be accounted for by the fact that the more generalized comparison of planet vs star is a more difficult question. Our preservice teachers, on the other hand, performed comparably to the primary school teachers studied by Summers and Mant [11], who scored 87% on an equivalent question. These results show that the relative size of planets and stars (with Earth vs the Sun being the most familiar example) is not widely known in middle school and hence must be explicitly addressed in class. Even at the tertiary level, instructors should be aware that not all students know this fundamental fact that planets are (far) smaller than typical stars.

We found no comparisons in literature for our second main error of not identifying the ozone layer as being far closer to the surface of Earth than any of the other objects in the list. More than 50% of middle school students, and more than 20% of preservice teachers thought the ozone layer to be further away than the center of Earth. Also, nontrivial numbers of middle school students thought that all of the other objects listed, except for the end of the Universe, were located closer to Earth's surface than the ozone layer (9%–21% preinstruction, 6%–12% postinstruction). It was unclear whether this reflected a belief that Earth's atmosphere actually encompasses these celestial bodies, or simply a lack of knowledge that the ozone layer is part of Earth's atmosphere in the first place. Another plausible explanation hinges on semantics: some students may be interpreting sky (as in "stars in the night sky") literally, failing to appreciate that sky in such a context may actually refer idiomatically to interplanetary or interstellar space, rather than something within Earth's atmosphere.

Our finding that more than 60% of middle school students and over a quarter of preservice teachers believed that the Pole star lies *within* our Solar System is less surprising given previous findings from literature. The nature and location of stars is the only topic beyond our Solar System investigated to any significant extent. Summers and Mant [11] found that only 77% of the 120 primary school teachers in their sample knew that the Sun is a star. Two separate questions probed whether the teachers believed that stars (plural) could be located within our Solar System, with 47% and 42%, respectively, responding positively to the question (while only 26% and 33%, respectively, correctly claimed the statements to be false, the remaining students responded "don't know" or did not answer). Trumper [36,37] probed Israeli junior high school students ($n = 448$, years 7–9) and senior high school students ($n = 378$, years 10–12) on the relative distance of the Moon, Pluto, and the stars from Earth in a MCQ with five alternatives. 36% and 49% of the student samples, respectively, responded correctly, whereas 51% and 41%,

respectively, placed the stars closer to Earth than Pluto (i.e., within our Solar System). In a similar question, Sadler [12] found that of 1414 U.S. high school students (years 8–12), 49% believed stars to be located closer than Pluto. These studies show a surprising and disconcerting level of agreement: in all groups, including our middle school students—ranging from junior high school students to primary school teachers—fewer than half know that stars (other than our Sun) are located well beyond our Solar System. Our preservice teachers performed somewhat better, but with 26% believing that the Pole star resides *within* our Solar System *after* instruction, the incorrect idea regarding the location of stars proves to be strongly held and highly resistant to change. Finally, Miller and Brewer [13] found that U.S. undergraduates overestimate the distance from Earth to the Moon, moderately underestimate the distance from Earth to the Sun, and dramatically underestimate the distances to the nearest star and to the nearest galaxy; the latter results echo our own findings in this study.

A few other results merit brief discussion. Our finding that a significant fraction of middle school students and a majority of preservice teachers interpreted solar system as stellar system is consistent with those from the South African IAQ study [29], where > 70% of the respondents made this same interpretation; this could be ascribed to the extensive coverage exoplanets have enjoyed in recent years in popular media. This familiarity with the concept (if nothing else) of exoplanets contrasts our more general finding of ignorance regarding the sizes, distances, and nature of basic astronomical entities. Also of note was our finding that the preservice teachers' scores on the different components of the aforementioned explaining task (i.e., scores for all individual objects) differed on average by only around 5 percentage points from those of the diverse sample of UCT students in the original IAQ study.

Our study offers some comparison between Norwegian preservice teachers and South African students in an Astro-101 course. It is interesting that these samples show remarkably similar results, given the stark contrast between the two countries in terms of socioeconomic and educational landscapes. Nevertheless, we do not consider these results particularly surprising, for two reasons. First, both tertiary samples were highly self-selected: the Norwegian preservice teacher sample comprised students who wished to become science teachers *and* who had been accepted into the largest institution for teacher training in the country; the South African students were a culturally diverse though self-selected group of students with an interest in astronomy (as the course was not compulsory) and who had been admitted to UCT, the highest ranked university on the African continent [38]. Consequently, neither sample is representative of its parent country as a whole, so the samples cannot be used for cross-cultural comparison.

B. Correlating students' qualitative and ranking knowledge

We found that for middle school students and preservice teachers alike, an ability to rank objects closely correlated with their knowledge of these objects. Students who displayed no qualitative knowledge of either star or planet did little better than chance in ranking them (54% correct), while students who correctly defined both objects in a pair were almost guaranteed to be able to rank them correctly (98% correct); those with partial knowledge fell somewhere in between.

This is a particularly interesting result, because it *suggests* that *if* students have satisfactory qualitative knowledge, they have a very high probability of possessing the corresponding ranking knowledge. However, the converse is *not* true: if students possess the correct ranking knowledge, its binary nature is too noisy to reveal any valuable information about a student's qualitative knowledge. The implication of this is that correctly ranking objects is *not* a reliable predictor for understanding astronomical phenomena; instead, students' qualitative knowledge is a better representation of their basic knowledge, and a better foundation for developing further understanding of astronomical phenomena. We note, however, that students' qualitative knowledge was probed through a single open question (focusing on the objects in the shorter ranking task), and that we did not study students' knowledge of the magnitudes of the sizes of or distances to the objects to be ranked. More research is also needed to investigate whether it is true in general that students who do not possess the correct ranking knowledge of a pair of objects do not have satisfactory qualitative knowledge of both objects in the pair.

Although middle school students made more mistakes in the ranking tasks, the patterns of correlation did not differ notably between the middle school and the preservice teacher samples, though there appears to be finer structures in the data that warrant further research attention. Consequently, the poorer performance of the middle school students on the ranking tasks may be attributable to poorer basic knowledge of astronomical objects in general.

VII. PEDAGOGICAL IMPLICATIONS AND FURTHER WORK

The implications of this study range from direct teaching recommendations—via more involved use of the ranking tasks as formative assessment for instructors, and using the ranking tasks for research on productive pedagogical interventions—to prompts for further exploration of the assumptions and propositions presented in the paper.

In the literature, knowledge of absolute size and distance is proposed as being important for understanding astronomical phenomena. However, we argue that it is *basic knowledge* (qualitative plus ranking knowledge) of astronomical objects that is crucial for understanding more complex phenomena. We consider it to represent a type of *gateway knowledge*: students risk not getting much out of further instruction if this basic knowledge is not in place [39]. While we acknowledge that this claim is yet to be supported by empirical evidence—which we consider a worthwhile avenue for further research—we nevertheless feel justified in advocating that educators ensure that students are explicitly taught what certain astronomical objects are, and do not assume that students arrive already equipped with this knowledge. It would be ill-considered to teach students about the detailed dynamics of the Milky Way galaxy, for example, if the students had not appreciated the constitutional fact that our galaxy contains hundreds of billions of stellar systems, one of which is our Solar System. Similarly, if students believe planets are larger than stars (as did more than 40% of our middle school students) or do not have a good idea of what a star is (two-thirds of our middle school students), it seems extremely implausible that they would be able to acquire an understanding of the formation history of the Solar System.

To evaluate whether students have sufficient basic knowledge to understand astronomical phenomena, the frequent use of formative assessments might be useful, for example, by using ranking tasks or writing brief definitions. Well-suited methods in this regard include Just-in-Time Teaching [40,41] and Peer Instruction [42,43].

Ranking tasks, in particular, can be used to gain important knowledge of students' ideas, and also of which pedagogical interventions have most impact on students' learning. Because of the simplicity of the ranking tasks, they can easily be completed at the beginning and end of an intervention, even if time is a significant constraint, or they can be integrated with other forms of evaluation. We also note that the analysis will be made even simpler if responses are captured directly on a computer.

To facilitate the study of students' knowledge of relative sizes and distances (which we term *ranking knowledge*) of common astronomical objects, we developed a new method for analyzing ranking tasks which enables detailed evaluation of a large number of students' conceptions of relative size and distance of many (5–10) objects. We believe that this can be a powerful tool both in teaching and research because it is cheap, simple for respondents to complete, and permits straightforward but detailed analysis. Our tool can also be extended to other disciplines that face similar challenges of scale, such as the microscopic worlds in physics and biology, or the enormous time scales of geology or cosmology.

Lelliott and Rollnick [5] (p. 1791) point out that size and distance are both "under-researched" and "undertaught." However, we believe that an important distinction has not yet been made in literature, viz. between relative rank and magnitude. The importance of rank is easy to argue: if one does not know that stars are significantly larger than planets, the basic construction of a solar system is unlikely

to make sense. However, the importance of magnitude is less clear. Beyond knowing that one object is larger than another, for what astronomical understanding is it important to know the details of how many orders of magnitude one object is more massive or further away than another? Consequently, we distinguish between ranking knowledge and scale knowledge, where the former merely reflects whether students know the ranking of objects in terms of size and distance, while the latter includes an understanding of magnitudes as well.

Based on our finding that *if* students have satisfactory qualitative knowledge of a pair of objects, *then* they are almost guaranteed to know how to rank them, we speculate that it is more important to focus on helping students learn *what* the objects are (i.e., qualitative knowledge) instead of focusing primarily on size and distance during instruction. From this finding, a powerful implication follows directly as an equivalent statement: viz., that if students *do not* possess the correct ranking knowledge of a pair of objects, then they almost certainly *do not* have satisfactory qualitative knowledge of both objects in the pair [44]. This means that the quantitative ranking task (which is simple to distribute and analyze, as opposed to the explaining task) might easily be used to evaluate the more important and extensive qualitative knowledge, indicating areas in need of instructional attention. It could also be used for similar purposes to evaluate interventions in research. The limit of the ranking task's usefulness is reached if students achieve a near perfect score, as this does not guarantee that students have satisfactory qualitative knowledge. However, the scores on ranking tasks were far from perfect in all of our samples (excepting perhaps the preservice teachers on the size task), and is thus unlikely to be a limitation in most samples of interest. Incidentally, it would also be of interest to investigate whether the correlations between qualitative and ranking knowledge is observed for objects in other disciplines.

In terms of direct implications for astronomy teaching, our analysis identified a number of prominent and persistent incorrect ideas students hold—middle school students to a greater degree than preservice teachers, but with the same trends present in either case—regarding the relative sizes of and distances to various astronomical objects. The most prevalent incorrect ideas all involved objects with which students have some personal experience, such as Earth, the Sun, and stars, but unfortunately the human experience of these objects is not conducive to developing an appropriate scale model of the universe. Without the introduction of scientific knowledge, one would not readily be able to deduce the correct relative sizes and distances of these objects; however, introducing the correct semantic (textbook) knowledge will not simply *override* one's experiential knowledge either. Rather, semantic and experiential knowledge must be reconciled, which requires explicit attention in instruction [13].

Recently, Yu *et al.* [45] argued that the planetarium can be a powerful tool for helping students visualize scale over many orders of magnitude; simulations and visualizations, more broadly, are powerful tools that can help students comprehend phenomena that play out on scales beyond normal sensory experience. Resnick *et al.* [46] showed that *hierarchical alignment* activities—in which students map increasingly larger scales to a familiar one using multiple analogies, locating the relative positions of all previous scales in each step—are beneficial for increasing the accuracy with which students can estimate temporal and spatial scales, even when those scales lie far beyond human perception. Nevertheless, the instructional interventions most effective at rectifying strongly held size and distance misconceptions remains an area in need of much further research attention. As our own results suggest, conventional astronomy instruction can in some contexts be effective (as with our preservice teachers), and in others completely ineffective (as with our middle school students) at remedying these misconceptions.

Work in AER has been conducted in a wide range of countries. In their review of articles published 1974–2008, Lelliott and Rollnick [5] reference studies from countries across the globe, including the USA, UK, Greece, Turkey, Israel, Estonia, India, China, Australia (including aboriginal children), and New Zealand (including Maori children). Although the authors do occasionally mention specific student alternative conceptions tied to cultural ideas (such as in India and aboriginal Australia), or an emphasis on the traditions of the educational system (such as rote learning in Estonia), the authors report no clear cultural differences in AER based on the literature they reviewed. A similar lack of different alternative conceptions in astronomy is reported in an earlier review article by Bailey and Slater [6]. This lack of cross-cultural difference suggests that the cognitive dimension of our shared, skewed experience may impede our basic knowledge of astronomical objects far more strongly than cultural ideas; this may explain the observed consistency in students' understanding of certain astronomical objects (stars, planets, galaxies, and so on) across national, linguistic, and cultural boundaries.

To conclude, we step back from the fine-grained details of the results we have been discussing and recall Carl Sagan's suggestion that there is no better demonstration of the "folly of human conceits"—including bloodthirst, greed, and our imagined self-importance—than the picture of Earth as a mere "mote of dust, suspended in a sunbeam… a very small stage in a vast cosmic arena." He argued that an appreciation of the scales of the cosmos—and by comparison, the minute scale of Earth—"underscores our responsibility to deal more kindly with one another, and to preserve and cherish […] the only home we've ever known" [47]. In a similar vein, Bertrand Russell argued that "philosophies [springing] from self-importance […] are best corrected by a little astronomy," yet that "the more we realize our

minuteness and our impotence in the face of cosmic forces, the more astonishing becomes what human beings have achieved" [48]. From this perspective, our findings would seem to suggest very great unfulfilled potential for using astronomy to influence positively the world view and general scientific thinking of all students [49,50].

Forthcoming papers in this series will focus on the analysis and results of the remaining questions in the NIAQ. From the longer-form questions given to the pre-service science teachers only, we study shifts in pedagogical behaviors preinstruction to postinstruction, and also probe further our assumption that students' basic descriptive knowledge is a better predictor of ability to explain astronomical phenomena than knowledge of size and distance. Separately, we consider a stark stratification along gender lines in the attitudes and performance of both the middle school students and the preservice teachers.

APPENDIX A: THE NIAQ IN ENGLISH

Each question (Q1–Q8) appeared on a separate page, so that the full NIAQ comprised 8 pages in total. Details about the formatting of questions appear in square brackets.

Q1a How *interesting* do you think *astronomy* is? Circle a number from 1 to 5, where 1 means "not at all interesting" and 5 means "extremely interesting."

Q1b How *important* to society do you think *astronomy* is? Circle a number from 1 to 5, where 1 means "not at all important" and 5 means "extremely important."

Q1c How much do you think *astronomers* have *not yet discovered*? Circle a number from 1 to 5, where 1 means "astronomers still have very much to discover" and 5 means "there is very little left for them to discover."

[Double line break]

Q1d How *interesting* do you think *physics* is? Circle a number from 1 to 5, where 1 means "not at all interesting" and 5 means "extremely interesting."

Q1e How *important* to society do you think *physics* is? Circle a number from 1 to 5, where 1 means "not at all important" and 5 means "extremely important."

Q1f How much do you think *physicists* have *not yet discovered*? Circle a number from 1 to 5, where 1 means "physicists still have very much to discover" and 5 means "there is very little left for them to discover."

Q2 A group of year-9 students is having an argument. Student A: "There is no difference between astronomy and astrology." Student B: "Nonsense! Astronomy and astrology are totally different." Student C: "They might be different but they are equally useful." Write down in detail what you will say to these students in order to settle their argument. [Blank box occupying three quarters of a page provided for response]

Q3 Two year-9 students are having an argument. Student A: "There is no life anywhere in the universe other than on Earth!" Student B: "I disagree. There has to be life elsewhere in the universe." With whom do you most closely agree? (Circle one letter) [Boxes containing letters A and B provided] Explain your choice in detail. [Blank box occupying half a page provided for response]

Q4 A group of year-9 students is having an argument. Student A: "I accept that the universe started with the Big Bang." Student B: "I don't believe in the Big Bang as its just a theory." Student C: "Alright, but then why do so many scientists seem to accept it?" Write down in detail what you will say to these students in order to settle their argument. [Blank box occupying three quarters of a page provided for response]

Q5a Rank the following from smallest to largest: *galaxy; planet; star; universe; solar system.* [5 numbered lines provided, with "smallest" in parenthesis next to number 1, and "largest" in parenthesis next to number 5.]

Q5b A year-9 student (who is interested in astronomy) asks you what is meant by each of the following: *galaxy; planet; star; universe; solar system.* Write *one sentence* per item. [2 ruled lines provided for each item]

Q6 A year-9 student asks you: "How do astronomers learn things about the universe? I know they primarily use telescopes, but what do the telescopes actually do?" Write down what you will tell this student. [Blank box occupying three quarters of a page provided for response]

Q7 Rank the following by their distance from the Earths surface: *centre of the Milky Way; edge of the observable universe; the asteroid belt; edge of the Solar System; the Moon; the Sun; the star Polaris; the ozone layer; centre of Earth; Neptune.* [10 numbered lines provided, with "closest" in parenthesis next to number 1, and "furthest" in parenthesis next to number 10.]

Q8 You overhear a group of year-9 students having a discussion. Student A: "I am always amazed that Earth just keeps going round and round the sun!" Student B: "Yes! How come it goes round and round and doesn't just fly off into space?" Student C: "Actually, I wonder the opposite: why doesn't Earth just spiral into the sun?" Explain to the

group, possibly with the aid of a diagram, how they should think about this (so that they don't lie awake at night getting nightmares, losing sleep, and ultimately failing their exams). [Blank box occupying three quarters of a page provided for response]

APPENDIX B: THE NIAQ IN NORWEGIAN

The formatting of the NIAQ in Norwegian was identical to the English version; only the text differed.

Q1a Hvor *interessant* synes du *astronomi* er? Ring rundt et tall fra 1 til 5, hvor 1 betyr "helt uinteressant" og 5 betyr "ekstremt interessant."

Q1b Hvor *viktig* synes du *astronomi* er for samfunnet? Ring rundt et tall fra 1 til 5, hvor 1 betyr "helt uviktig" og 5 betyr "ekstremt viktig."

Q1c Hvor mye tror du *astronomer ennå ikke har oppdaget*? Ring rundt et tall fra 1 til 5, hvor 1 betyr "astronomer har svært lite igjen å oppdage" og 5 betyr "astronomer har fortsatt svært mye igjen å oppdage."

Q1d Hvor *interessant* synes du *fysikk* er? Ring rundt et tall fra 1 til 5, hvor 1 betyr "helt uinteressant" og 5 betyr "ekstremt interessant."

Q1e Hvor *viktig* synes du *fysikk* er for samfunnet? Ring rundt et tall fra 1 til 5, hvor 1 betyr "helt uviktig" og 5 betyr "ekstremt viktig."

Q1f Hvor mye tror du *fysikere ennå ikke har oppdaget*? Ring rundt et tall fra 1 til 5, hvor 1 betyr "fysikere har svært lite igjen å oppdage" og 5 betyr "fysikere har fortsatt svært mye igjen å oppdage."

Q2 En gruppe 10.-klassinger har en diskusjon. Elev A: "Det er ingen forskjell mellom astronomi og astrologi." Elev B: "Pølsevev! Astronomi og astrologi er helt forskjellige." Elev C: "De er kanskje forskjellige, men de er begge like nyttige." Skriv i detalj hva du ville sagt til disse elevene for å avklare diskusjonen.

Q3 To 10.-klassinger har en diskusjon. Elev A: "Det finnes ikke liv noe annet sted i universet enn på jorda!" Elev B: "Ikke enig. Det må finnes liv andre steder i universet." Hvem er du mest enig med? Ring rundt én bokstav. […] Gi en detaljert begrunnelse for ditt valg.

Q4 En gruppe 10.-klassinger har en diskusjon. Elev A: "Jeg godtar at universet startet med Big Bang." Elev B: "Jeg tror ikke på Big Bang, fordi det bare er en teori." Elev C: "Greit, men hvorfor er det da så mange forskere som godtar den?" Skriv i detalj hva du ville sagt til disse elevene for å avklare diskusjonen.

Q5a Rangér følgende fra minst til størst: *galakse; planet; stjerne; univers; solsystem..* [(Minst)…(Størst)]

Q5b En 10.-klassing (som er interessert i astronomi) spør deg hva som menes med følgende; skriv *én setning* for hver: *galakse; planet; stjerne; univers; solsystem.*

Q6 En 10.-klassing spør deg: "Hvordan lærer astronomer ting om Universet? Jeg vet de hovedsakelig bruker teleskoper, men hva gjør egentlig teleskopene?" Skriv hva du ville fortalt denne eleven.

Q7 Rangér følgende etter deres avstand fra Jordas overflate: *sentrum av Melkeveien; enden av det observerbare univers; asteroidebeltet; enden av solsystemet; månen; solen; Polarstjernen; ozonlaget; jordas sentrum; Neptun.* [(Nærmest)…(Lengst unna)]

Q8 Du overhører en gruppe 10.-klassinger som har en diskusjon. Elev A: "Jeg blir alltid like forbløffet over at jorda fortsetter å gå rundt og rundt sola!" Elev B: "Ja! Hvorfor går den rundt og rundt uten å fly av gårde ut i verdensrommet?" Elev C: "Jeg lurer faktisk på det motsatte: hvorfor faller ikke jorda i en spiral inn mot sola?" Forklar gruppen, kanskje ved hjelp av et diagram, hvordan de burde tenke på dette (så de ikke ligger våkne om nettene, får mareritt, mister søvn og til slutt gjør det dårlig på eksamen).

---

[1] R. R. Hake, Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses, Am. J. Phys. **66**, 64 (1998).

[2] L. C. McDermott and E. F. Redish, Resource letter: Per-1: Physics education research, Am. J. Phys. **67**, 755 (1999).

[3] A. A. diSessa and P. Cobb, Ontological innovation and the role of theory in design experiments, J. Learn. Sci. **13**, 77 (2004).

[4] U. Eriksson, C. Linder, J. Airey, and A. Redfors, Who needs 3D when the universe is flat?, Sci. Educ. **98**, 412 (2014).

[5] A. Lelliott and M. Rollnick, Big Ideas: A review of astronomy education research 1974-2008, Int. J. Sci. Educ. **32**, 1771 (2010).

[6] J. M. Bailey and T. F. Slater, A review of astronomy education research, Astron. Educ. Rev. **2**, 20 (2004).

[7] C. S. Wallace, E. E. Prather, and D. K. Duncan, A study of general education astronomy students' understandings of cosmology. Part IV. Common difficulties students experience with cosmology, Astron. Educ. Rev. **11**, 010104 (2012).

[8] J. M. Bailey and D. Lombardi, Blazing the trail for astronomy education research, J. Astron. Earth Sci. Educ. **2**, 77 (2015).

[9] C. Bakas and T. A. Mikropoulos, Design of virtual environments for the comprehension of planetary phenomena based on students' ideas, Int. J. Sci. Educ. **25**, 949 (2003).

[10] M. Cin, Alternative views of the solar system among Turkish students, Int. Rev. Educ. **53**, 39 (2007).

[11] M. Summers and J. Mant, A misconceived view of subject-matter knowledge in primary science education: a response to golby *et al.* 'some researchers' understanding of primary teaching, Res. Papers Educ. **10,** 303 (1995).

[12] P. M. Sadler, The initial knowledge state of high school astronomy students, Ph.D. thesis, Harvard University, Cambridge, MA, 1992.

[13] B. W. Miller and W. F. Brewer, Misconceptions of astronomical distances, Int. J. Sci. Educ. **32,** 1549 (2010).

[14] P. M. Sadler, Psychometric models of student conceptions in science: Reconciling qualitative studies and distractor-driven assessment instruments, J. Res. Sci. Teach. **35,** 265 (1998).

[15] T. M. Fanetti, The relationships of scale concepts on college age students' misconceptions about the cause of lunar phases, Master's thesis, Iowa State University, Ames, Iowa, 2001.

[16] Central Intelligence Agency, CIA World Factbook—Norway, (2016), online: accessed 22 November 2016.

[17] United Nations Development Programme, 2015 Human Development Report, (2015), online: accessed 22 November 2016.

[18] Legatum, The Legatum Prosperity Index, (2016), online: accessed 22 November 2016.

[19] Wikipedia, Norway—Wikipedia, The Free Encyclopedia, (2016), online: accessed 22 November 2016.

[20] Store norske leksikon: Skole og utdanning i norge, (2015), online: accessed 30 June 2015.

[21] P. R. Ekeland, O.-I. Johansen, O. Rygh, and S. B. Strand, *Tellus 8: Naturfag for ungdomstrinnet* (Aschehoug, Oslo, 2006).

[22] M. Hannisdal, J. Haugan, J. Nyberg, and M. Frøyland, *Eureka! 8, Naturfag for ungdomstrinnet, grunnbok* (Gyldendal, Oslo, 2006).

[23] H. S. Finstad, J. Kolderup, and E. C. Jørgensen, , *Trigger 8: Naturfag for ungdomstrinnet* (Cappelen Damm, Oslo, 2006).

[24] T. L. O'Kuma, D. P. Maloney, and C. J. Hieggelke, *Ranking Task Exercises in Physics* (Prentice Hall Upper Saddle River, NJ, 2000), Vol. 26.

[25] D. P. Maloney, Charged poles?, Phys. Educ. **20,** 310 (1985).

[26] S. J. Slater, P. S. Sharon, and D. J. Stork, Analysis of individual test of astronomy standards (toast) item responses, J. Astron. Earth Sci. Educ. **2,** 89 (2015).

[27] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms* (MIT Press, Cambridge, MA, 2009).

[28] In the future, it might also be helpful to include a simple, nonastronomical ranking task (e.g., rank a house, an apple, a mountain, etc., from smallest to largest) to identify specific students who have difficulty interpreting the ranking tasks as indented.

[29] V. Rajpaul, S. Allie, and S.-L. Blyth, Introductory astronomy course at the University of Cape Town: Probing student perspectives, Phys. Rev. ST Phys. Educ. Res. **10,** 020126 (2014).

[30] C. Lindstrøm, V. Rajpaul, M. Brendehaug, and M. C. Engel, Perspectives on astronomy: probing Norwegian pre-service

teachers and middle school students, in *Proceedings of the Physics Education Conference, College Park, MD, July 29–30, 2015,* edited by A. D. Churukian, D. L. Jones, and L. Ding (AIP, New York, 2016).

[31] D. W. Hudgins, E. E. Prather, D. J. Grayson, and D. P. Smits, Effectiveness of collaborative ranking tasks on student understanding of key astronomy concepts, Astron. Educ. Rev. **5,** 1 (2006).

[32] E. Jones, T. Oliphant, and P. Peterson, SciPy: Open source scientific tools for Python (2001).

[33] Admittedly stellar remnants such as white dwarf and neutron stars are smaller than most planets, but we have no reason to believe students adopted this extremely narrow interpretation of the term "star."

[34] When comparing Q7 and Q5b scores for the preservice teachers, the χ^2 test did not allow us to *reject* the null hypothesis of complete statistical independence: this is *not* the same as proving there was definitely no form of association between the scores. Thus, the results from the χ^2 tests and linear-correlation analysis are not incompatible.

[35] V. R. welcomes correspondence from anyone who would like to discuss such software, perhaps with a view to applying it in their own research.

[36] R. Trumper, A cross-age study of junior high school students' conceptions of basic astronomy concepts, Int. J. Sci. Educ. **23,** 1111 (2001).

[37] R. Trumper, A cross-college age study of science and nonscience students' conceptions of basic astronomy concepts in preservice training for high-school teachers, J. Sci. Educ. Technol. **10,** 189 (2001).

[38] Times Higher Education, World University Rankings, (2018), online: accessed 1 March 2018.

[39] Whereas we believe basic knowledge to be sufficient for understanding most astronomical phenomena taught in schools and even Astro-101, it is unlikely to be sufficient at higher levels of education. One could propose the concept of *advanced knowledge* of objects, which would include and extends basic knowledge; it might comprise *technical knowledge* (which extends qualitative knowledge) and *scale knowledge* (which refers to the relative rank *and* magnitude of a pair of objects). We defer consideration of such possibilities to future work.

[40] G. M. Novak, E. T. Patterson, A. D. Gavrin, W. Christian, and K. Forinash, Just in time teaching, Am. J. Phys. **67,** 937 (1999).

[41] G. M. Novak, Just-in-time teaching, New Directions Teaching Learning **2011,** 63 (2011).

[42] E. Mazur, Peer instruction: Getting students to think in class, AIP Conf. Proc. **399,** 981 (1997).

[43] E. Mazur and M. D. Somers, Peer instruction: A user's manual, Am. J. Phys. **67,** 359 (1999).

[44] Using formal propositional logic, note that $(A \rightarrow B) \Leftrightarrow (\neg B \rightarrow \neg A)$. On the other hand, $(A \rightarrow B) \not\Leftrightarrow (B \rightarrow A)$.

[45] K. C. Yu, K. Sahami, and J. Dove, Learning about the scale of the solar system using digital planetarium visualizations, Am. J. Phys. **85,** 550 (2017).

[46] I. Resnick, N. S. Newcombe, and T. F. Shipley, Dealing with big numbers: Representation and understanding of

magnitudes outside of human experience, Cogn. Sci. **41,** 1020 (2017).

[47] C. Sagan, *Pale Blue Dot: A Vision of the Human Future in Space* (Random House, New York, 1994).

[48] B. Russell, *The Basic Writings of Bertrand Russell*, edited by R. E. Egner and L. E. Denonn (Routledge, London, 2009).

[49] M. R. Matthews, Science, worldviews and education: An introduction, Sci. Educ. **18,** 641 (2009).

[50] C. S. Wallace, E. E. Prather, and B. M. Mendelsohn, Astro 101 students' perceptions of science: Results from the thinking about science survey instrument, Astron. Educ. Rev. **12,** 010101 (2013).