A LIVELY ELECTRONIC COMPENDIUM OF RESEARCH, NEWS, RESOURCES, AND OPINION

Astronomy Education Review

2013, AER, 12(1), 010102, http://dx.doi.org/10.3847/AER2012038

Investigating Student Ideas about Cosmology I: Distances and Structure

Kim Coble
Chicago State University, Chicago, Illinois 60628
Carmen T. Camarillo
Chicago State University, Chicago, Illinois 60628
Laura E. Trouille
Northwestern University, Evanston, Illinois 60201
Janelle M. Bailey
University of Nevada, Las Vegas, Las Vegas, Nevada 89154
Geraldine L. Cochran
Florida International University, Miami, Florida 33199
Melissa D. Nickerson
Chicago State University, Chicago, Illinois 60628
Lynn R. Cominsky
Sonoma State University, Rohnert Park, California 94928
Received: 09/06/12, Revised: 12/03/12, Accepted: 02/08/13, Published: 06/03/13

Abstract

Recently, powerful new observations and advances in computation and visualization have led to a revolution in our understanding of the structure of the Universe. As the field of cosmology advances, it is of interest to study how student ideas relate to scientific understanding. In this paper, we examine in-depth undergraduate students' ideas on distances and structure in the Universe as students progress through a general education astronomy integrated lecture and laboratory course with a focus on active learning. The study was conducted over five semesters at an urban, minority-serving institution. The data collected include individual interviews ($N = 15$) and course artifacts ($N \sim 60$), such as precourse homework essays, prelab surveys, and midterm and final exam questions in a variety of formats. We find that students are fairly successful at tasks involving relative distances, but struggle with absolute distances; have difficulty going beyond an elementary model of the Solar System as the Sun and planets; struggle to visualize galactic halos; but successfully increase their understanding of the hierarchical nature of structure in the Universe throughout the semester.

1. INTRODUCTION

Cosmology is a field of study that is rapidly advancing our current knowledge of the Universe. Gains in cosmological science have been vast, but their impact on education has been limited. It is important to bridge this new knowledge of cosmology to the classroom, because understanding the underpinnings of the Universe can deepen students' sense of wonder and help them appreciate their origins, in the broadest sense. Cosmology also provides a powerful means to help students understand the link between a scientific worldview and the data upon which it is based.

Undergraduate astronomy courses have had difficulty staying current with rapidly unfolding cosmological knowledge. Some members of the education community (e.g., Pasachoff 2002) have called for increased use of modern topics and research results in these courses. With calls for science education reform at all levels (American Association for the Advancement of Science [AAAS] 1990, 1993; Bransford *et al*. 1999; Fox and

Hackerman 2003; National Research Council [NRC] 1996, 2003, 2012), and the presence of cosmological concepts in both the K-12 *National Science Education Standards* (NRC 1996) and the *Project 2061 Benchmarks for Science Literacy* (AAAS 1993), it is imperative that we design effective instruction to counteract student misconceptions, build upon correct ideas, and provide scaffolds for new understandings at the college level as well (Donovan and Bransford 2005).

As revealed by Deming and Hufnagel (2001), most beginning astronomy undergraduate (a.k.a. ASTRO 101) students have never had an astronomy course of *any* kind. Thus, their understanding of the Universe has been formed through exposure in the media (news, television, movies) or in their limited astronomy experiences in K-12 education. Additionally, most ASTRO 101 students will not take any further science classes (Rudolph *et al.* 2010). Thus, whatever understanding they gain from ASTRO 101 will have to serve them as they take their places as adult citizens and perhaps even as teachers. To support the development of a more scientifically literate society, we need to provide ASTRO 101 courses based on a modern view of the Universe and a more complete treatment of science (e.g., a treatment in which students work with real astronomical data to come to new understandings).

Decades of science education research has shown that students' ideas do not always match scientific understanding of concepts, and that this is true for all grade levels and student backgrounds (see, e.g., Duit 2006 and references therein). Determining the range and frequency of these "alternative conceptions" is an important first step to improving instructional effectiveness in cosmology, and astronomy education researchers are beginning to document them. For example, Prather, Slater, and Offerdahl (2002) examined middle school, high school, and college students' ideas about the Big Bang. A subsequent study by Wallace, Prather, and Duncan (2012) included other cosmological topics such as the expansion of the Universe, reading Hubble diagrams, and dark matter. Other alternative conceptions about the age of the Universe, Earth's position in it, and its history are documented using the Astronomy Diagnostic Test (Deming 2002). Miller and Brewer (2010) discuss students' ideas of absolute astronomical scales. Work from nanoscience and geoscience education and on size and scale more generally (e.g., Cheek 2012; Delgado *et al.* 2007; Tretter, Jones, and Minogue 2006) can also inform our work on astronomical sizes and scales. Additional information about general astronomy alternative conceptions is included in reviews by Bailey and Slater (2003) and Lelliott and Rollnick (2010). We leave a discussion of the most relevant results from these papers to the following sections.

Our group is bringing new tools and advances to cosmology instruction through research on undergraduate learning in cosmology as well as through the development of a series of web-based cosmology learning modules for general education undergraduate students (Coble *et al.* 2012; Coble *et al.* in preparation). We have structured both our research and our curriculum development around three major cosmological themes: (1) structure—the vast distances, timescales, and hierarchical nature of structure; (2) composition—the Universe is composed of not just regular matter but also dark matter and dark energy; and (3) change—the Universe is dynamic and evolving, exemplified by the Big Bang model and age and expansion of the Universe. Integral to our approach is including how this knowledge is supported by observational and experimental evidence and why these processes occur (according to the laws of physics).

The present paper is one of a series examining the nature and frequency of students' ideas about structure, composition, and change. Our aim is to thoroughly catalog students' ideas, from a single institution, using multiple data sources. We are interested in documenting not only students' preinstructional ideas as other studies have done but also the range of student ideas as sampled over the course of the semester. We use a mixed-methods approach including both qualitative and quantitative analysis to derive a bigger picture and deeper understanding of students' ideas. A mixed-methods approach using multiple data streams can be a powerful approach to research questions, allowing for comparisons and providing a rich, flexible data set (Beichner 2009; Kregenow, Rogers, and Constas 2010).

In this article (Paper I), we examine student ideas on distance and structure. Using a similar methodology, in Coble *et al.* (submitted, Paper II), we examine student ideas on the composition of the Universe. In Trouille *et al.* (submitted, Paper III), we examine student ideas on the age, expansion, Big Bang, and history of the Universe. This approach complements that of Bailey *et al.* (2012), where we presented the results of a nationwide, open-response, precourse survey on various cosmological topics, including distances, structure, composition, Big Bang, and age of the Universe. We have also collected information about the geometry, accelerating expansion, and fate of our Universe, which will form the basis of future analyses.

In Section 2 of this paper, we describe our methodology, including the setting, participants, data sources, and analysis procedure. We then discuss students' ideas on distances and structure in Sections 3 and 4, respectively.

These data are presented chronologically as they were sampled throughout the semester, alternating results and implications for each subtopic. In Section 5, we conclude with a discussion of our most important results.

2. METHODS

2.1 Setting and Participants

Over five semesters (Fall 2008, Spring 2009, Spring 2010, Fall 2010, and Spring 2011), multiple sources of data were collected from the general education astronomy course at Chicago State University (CSU), an urban, minority-serving institution. The demographics of the students in the astronomy classes are representative of the university's undergraduates as a whole (84% African-American, 7% Latino, 71% women, median age ∼25; The Office of Institutional Effectiveness and Research, 2011).

The course is an integrated lecture and laboratory class, with approximately 15 students per semester. It covers the major topics typically taught in an ASTRO 101 course (Slater *et al*. 2001), with somewhat more of an emphasis on cosmological topics than is typical at other institutions. The class meets four hours per week for 15 weeks. Coble taught the course all semesters except one; Trouille taught the course during the Spring 2011 semester, using Coble's curricular materials.

As guiding principles for the class, we want students to learn both the content and processes of science, including making predictions and testing them experimentally, asking questions in order to gain understanding, relating science to everyday life, and reflecting on results; the class forms a scientific community. Ours is an active classroom; interactive lectures are integrated with short and long tasks, such as CSU-developed worksheets, *Lecture-Tutorials* (Adams *et al*. 2005; see Note-1), activities from *Mastering Astronomy* (see Note-2), and longer laboratory-oriented activities. Laboratory activities were developed from existing verification-style (or "cookbook") laboratories following the guiding principles for our course materials. Students also complete an observing project using the Global Telescope Network (see Note-3). A CSU-developed course workbook ties materials together. We will provide more details on the activities relevant to specific cosmological topics as students' ideas are addressed later in this paper.

The course schedule, as well as a list of when various data were collected, is given in Table 2.1. Weekly topics include material covered in lecture, laboratory, and other activities. The schedule was the same for all five semesters of data collection. Interviews were conducted over four semesters.

2.2 Data Collection

The data consist of course artifacts: precourse homework essays ($N = 55$), laboratory pretests (a.k.a. prelab surveys, N ∼ 60), and exam responses ($N \sim 60$), as well as in-depth interviews ($N = 15$). Course artifacts were collected over five semesters and interview data were collected over the final four semesters. Each of these data types is described further below. These N's, as well as those reported throughout the paper, are totaled over all semesters unless otherwise indicated. The number of responses can differ across questions for a variety of reasons, including: not every exam question was asked every semester, not every student turned in every assignment or was present in class for all prelab surveys, and there was some attrition over the course of each semester. The number of responses for each question will be presented in the appropriate data tables in each following section.

During the first week of class of each semester, students were assigned to write a 2-3-page homework essay describing the Universe as they currently understood it. Students were urged to describe what they really thought and were graded on completeness only, not the accuracy of their work. Students were asked to address their ideas and beliefs about three themes: (1) the physical size and structure of the Universe, (2) how the Universe changes over time, and (3) how humans fit into the big picture. These themes were framed with guiding questions, but students were not required to respond to every one. The section of the assignment relevant to distance and structure prompted students as follows:

---

[1] Although a newer edition is now available, the first edition was used in our course at the time of data collection.

[2] masteringastronomy.com.

[3] gtn.sonoma.edu.

Table 2.1. Relative schedule of topics, cosmology-related laboratory activities, and data collection points. The schedule was the same for all five semesters of data collection. Interviews with students (labeled A - O) were collected over four semesters.

Weekly topics	Laboratory[a]	Assessment/interview[b]
Introduction; Scale of the Universe		
Scale of the Universe; Process of science; History of astronomy	Laboratory #1: Scales of the Universe ($N = 74$) for pretest	HW #1 due ($N = 55$)
Looking at the sky; Seasons		Laboratory #1 due
Moon phases; Motion, gravity, energy		
Motion, gravity, energy; Light		
Light and telescopes		Exam #1 ($N_{max} = 65$)[c]
Solar system: exploration, formation, climate change, exoplanets		
The Sun; Stars: lifetimes, properties, classification		Interview: A
Stellar evolution; our Galaxy		Interview: B
Other galaxies; Dark matter	Laboratory #8: Mass of Galaxies ($N = 47$) for pretest	Interview: C
Measuring distances	Laboratory #9: Measuring Distances ($N = 36$) for pretest	Laboratory #8 due Interviews: D, E
Expansion and age of the Universe	Laboratory #10: Hubble Law	Laboratory # 9 due Interview: F
Big Bang: history of Universe, fate of Universe		Laboratory #10 due Interview: G
Observing project review panel	“Galaxy challenge”	Interviews: H, I
Life in the Universe		Exam #3 ($N_{max} = 56$); Interviews: J, K, L
Present observing projects, review		Interviews: M, N
Final Exam		Final Exam ($N_{max} = 58$) Interview: O

[a] Laboratories #2-7 do not relate to cosmology topics and so are not listed here.

[b] Assessments were given before the weeks’ topics were covered in class.

[c] N_{max} is the maximum number of responses to questions on a given exam. The number of responses might have been less for specific questions because not every question was asked every semester. The number of responses for each question is presented in the relevant data tables.

> “Describe what you think the physical size and structure of the Universe is like. (For example, imagine looking around while floating away from Earth, out into space. What kinds of things will you see as you drift farther and farther away from home? How are they arranged? Are stars and galaxies spread uniformly throughout space? What kinds of patterns or structures do you notice? How much empty space do you see? How big is the Universe? How long does it take to travel to various destinations? Does the Universe have an edge somewhere, or does it go on forever?)”

These precourse homework essays provide us with information on students’ preinstruction ideas. Because students were not required to specifically address each guiding question, but rather just the major themes, all response numbers obtained for the precourse homework essays should be taken as lower limits (i.e., it is possible that more students hold a given idea but did not discuss it specifically in their essays).

Other class data included prelab surveys, or “pretests,” and midterm and final exams. Laboratory pretests were open-ended written response format, administered after lecture but before laboratory activities. Exam questions included long-format open-response (essay) questions and short-format questions taken from various sources, such as textbook question banks and questions created by our group or other ASTRO 101 instructors who have shared their materials with us. Short-format exam questions included matching and ranking questions, multiple-choice (MC), true-false (T/F), and fill-in-the-blank (FIB) items.

The interviews were semistructured, based on the precourse survey questions described in Bailey *et al*. (2012). Interviews lasted approximately 30-40 min each and were transcribed afterward. Semi-structured interviews are used when the researcher anticipates that questions will require discussion and possibly follow-up questions (Rubin and Rubin 2005). Not all interviewees were asked all of the main questions. The purpose of a semi-structured interview is to allow the interviewee's responses to guide the interview and to use questions to get "a conversation going on a subject and ensure that the overall subject is covered" (Rubin and Rubin 2005, p. 13). In our tables and figures, we note the number of interviewees who were asked each question. Interviews took place throughout the latter half of each of four semesters. Thus, we are able to examine student ideas throughout the learning process, including postinstruction for a subset of the topics. We use quotes and themes from the interviews as illustrative examples of the student ideas we gathered through our course artifacts.

2.3 Data Analysis Procedure

We used a mixed methods approach in analyzing the data. For the precourse essays, pretests, exam essay questions, and interviews, we carried out an iterative process of thematic coding to generate a comprehensive list of themes. We then identified the fraction of students who discussed a given theme in their response. Open-response essays from the pretests and exams were also coded for degree of completeness and correctness, by comparing the actual response to the desired response (which may contain more than one element). The rubric for this analysis is detailed in Table 2.2. Correct answers are provided with the questions in the appropriate data analysis tables in the Appendix.

We used the Kruskal-Wallis (KW) test to determine whether we could aggregate results from different semesters. One advantage of the KW test is that it is applicable to data sets in which the number of values from each semester can be of equal or unequal lengths (i.e., it is valid even if there are different numbers of students in each semester). The KW test is a nonparametric method for testing the hypothesis that three or more sample populations (in our case, results from each semester) have the same mean distribution, against the hypothesis that they differ (see Note-4). Here, we used a conservative significance level of 0.1. If our p-value is greater than 0.1, we do not reject the null hypothesis that the semester results come from the same parent population. In other words, $p > 0.1$ means we can use the aggregated results because the semesters do not appear to differ significantly from one another. For each question, we ran a KW test to determine whether the different semesters' data could be combined. We used a Kolmogorov-Smirnov (KS) test for questions where there were only two semesters to combine. The KS test is also a nonparametric method and we used the same conservative significance level of 0.1 to determine whether the two semesters' results could be aggregated. In every case, the p-value was greater than 0.1 (ranging from 0.15 to 0.9), so we felt comfortable aggregating results across semesters.

In order to not disrupt the flow of the analysis, we provide summary tables of the aggregated data in the main body of the paper. In corresponding tables in the Appendix, we provide the full text of the questions asked, the

Table 2.2. Coding scheme for open-response prelab and exam essay questions.

Code	Meaning	Description
C	Correct	The response was complete and contained no wrong statements
I	Incomplete	The response was missing one or more of the identified elements required for a correct answer
P	Partial	The response contained both incorrect and correct elements
W	Wrong	No element of the response matched the identified elements of a correct answer
T	True but irrelevant	Included statements that were true but did not address the question in any meaningful way
NS	Non-scientific	Nonscientific response
NR	No response	No response

---

[4] The parametric equivalent to the KW test is the one-way analysis of variance (ANOVA). The KW test is an extension of the Mann-Whitney U test (which analyzes sample pairs for differences) to three or more groups.

correct answers used for coding, and detailed results for each item broken down by semester, as well as the KW H-statistic and p-value or the KS-statistic and the p-value as appropriate.

In each following section, the results of the course artifacts are presented in chronological order (precourse homework essays, laboratory pretests, midterm exams, and final exams) followed by interview data. This creates a coherent narrative of students' ideas over the course of the learning process. This sometimes limits the physical proximity of matched data but better elucidates what students are thinking as they acquire new knowledge. Some description of the curriculum is given as context for the environment in which the development of student ideas is taking place. However, discussions of the curriculum will be kept brief, as it is not our intention in this study to measure its effectiveness.

3. DISTANCES

In this section, we will discuss students' ideas regarding astronomical sizes and distances, including: the meaning of the light-year and implications for space travel and lookback time; absolute and relative size and distance scales; and techniques for distance measurement, such as parallax and the inverse square law/standard candles. Understanding astronomical distances and sizes is one of the most fundamental insights a student can gain from an astronomy course. Furthermore, the concept of size and scale is a unifying theme in national science education reform documents (AAAS 1993; NRC 1996) across STEM disciplines. Astronomers are also continually refining measurements of distances in the Universe, with significant results emerging within the past 15 years or so (e.g., Hubble Space Telescope Key Project results of Freedman *et al.* 2001, which draws from several methods of distance measurement).

At CSU, students are initially introduced to objects in the Universe and their sizes and distances during the first two weeks of class. After a brief lecture, students begin by working in groups to rank objects by size and distance using the *Cosmic Survey* activity (Smithsonian Astrophysical Observatory, 2003) to elicit their ideas. They build on this with required readings, further lectures on size and scale, and ranking tasks from *Mastering Astronomy* both in class and for homework (for other examples of ranking tasks in astronomy and physics see, e.g., Hudgins *et al.* 2006 and O'Kuma, Maloney, and Hieggelke 2000). Students learn the meaning of the term "light-year" and the relationships between distance, speed, and time through lecture and the *Lecture-Tutorial* "Looking at Distant Objects." Students cap off their introductory work on sizes and distances with Laboratory 1: *Scales of the Universe*, which uses the *Mastering Astronomy* tutorial of the same name. We have written a wrapper for the laboratory, such that students must make predictions about what they will discover in each section of the tutorial before proceeding with the results. The theme of distances is reprised midsemester when students do the whole-class activity *Toilet Paper Solar System*. In this classic activity, where 1 toilet paper square = 1 AU, students create a scale model of the Solar System as a class (see Note-5). They work more with distances in our Galaxy using the *Lecture-Tutorial* "Milky Way Scales." Near the end of the semester, students learn parallax, standard ruler, and standard candle distance measurement techniques through readings, lecture, and Laboratory 9: *Measuring Cosmic Distances*.

3.1 Light-Years and Space Travel

3.1.1 Results

We looked for emergent themes regarding student conceptions of the term light-year and space travel, first analyzing the 55 precourse homework essays. Students were not prompted specifically to mention speed of light or the term light-year; however, they were asked how long it would take to travel to various destinations.

The precourse homework essays were analyzed for students' ideas regarding space travel, split both by type of objects they mentioned and whether the travel time that they quoted was reasonable or unreasonable (Figure 3.1). Since students did not specify the speed or mode of transportation in any instances, we define "unreasonable" to be anything that would exceed the speed of light.

In examining the precourse homework essays for use of the term light-year, we found that nine students mentioned the term. Four students used light-year correctly as a measurement of distance and five students used light-year incorrectly as a measurement of time (Figure 3.2).

---

[5] As an example of this type of activity, see http://solar.physics.montana.edu/tslater/plunger/tissue.htm. Our model also included the beginning and ending radii of the asteroid belt and Kuiper belt, as well as the dwarf planets Haumea, Makemake, and Eris.

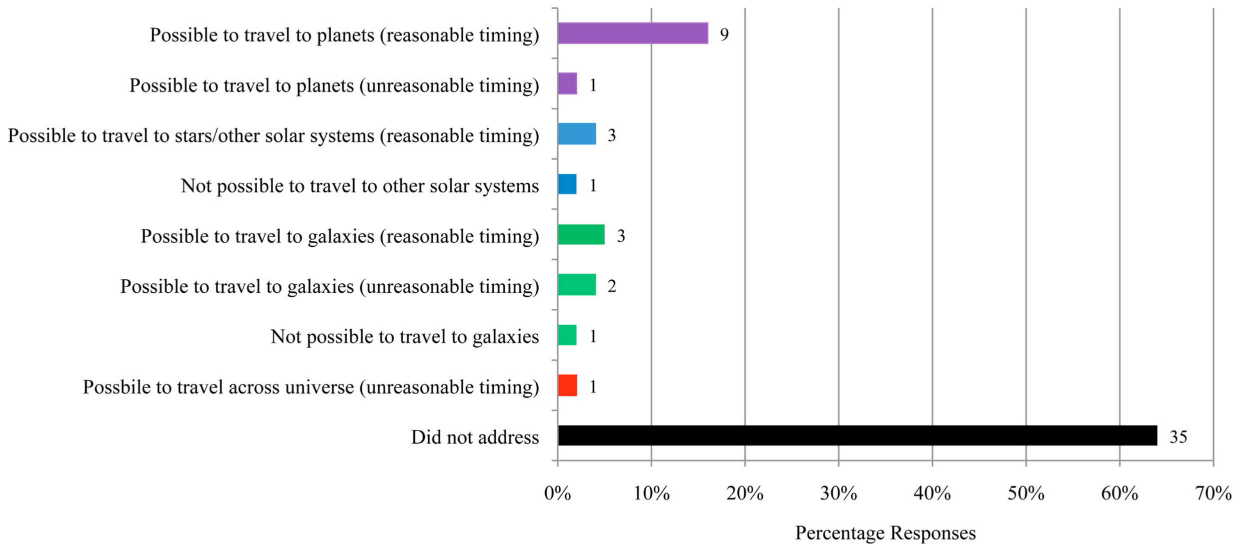


Figure 3.1. Precourse homework essays: Space travel. $N = 55$ essays were collected and analyzed for themes. It is possible for an essay to be coded with more than one theme. Since essays were free-form, the number shown for each category reflects a lower limit on the number of students who hold a particular concept.

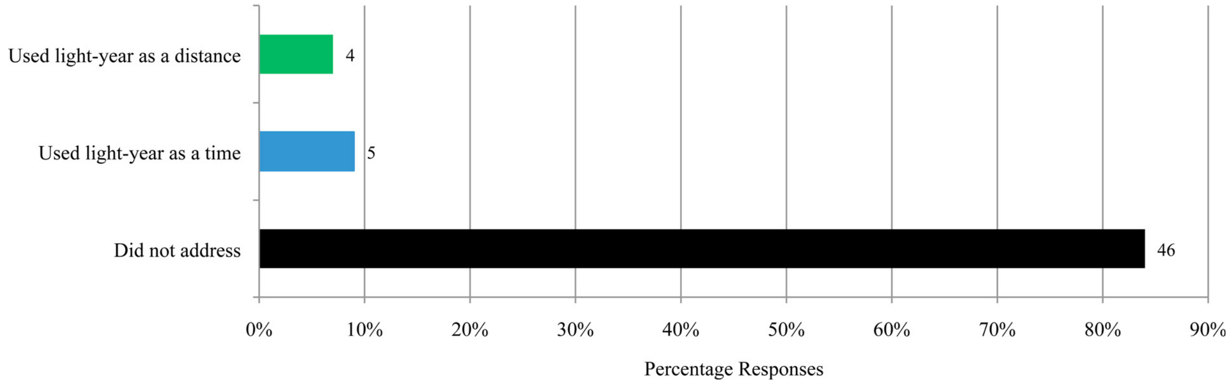


Figure 3.2. Precourse homework essays: Speed of light and the light-year. $N = 55$ essays were collected and analyzed for themes. It is possible for an essay to be coded with more than one theme. Since essays were free-form, the number shown for each category reflects a lower limit on the number of students who hold a particular concept.

Tables 3.1, and A3.1 (in the Appendix) show the results of the pretest for Laboratory 1: *Scales of the Universe* (administered after lecture but before students attempted the laboratory). The laboratory pretest examines student ideas regarding light-years and space travel. In the first part of a two-part question, students were asked to determine whether or not they had enough time to travel to the star Sirius within a given time period (three weeks) at half of the speed of light and explain how long it would take. According to the coded responses ($N = 74$), only a small fraction of students were Correct (16%). Responses were coded as Correct if students answered “no” it is not possible and provided the correct explanation (which involved multiplying the distance to the star in light-years by a factor of two to get the travel time in years). The most common reason for a code of Incomplete was that the student only answered “no” for the first part of the question and either did not provide an

Table 3.1. Laboratory 1 pretest: Half speed of light.

N	C	I	P	W	T	NS	NR
74	**16%**	23%	24%	28%	7%	0%	1%

Table 3.2. Laboratory 1 pretest: Speed and scale.

Question part	*N*	C	I	P	W	T	NS	NR
a. Content	74	**46**%	3%	27%	14%	3%	0%	7%
b. Reasoning	74	**5**%	14%	0%	3%	4%	0%	67%

explanation or repeated the question in their response. Several Partials were given to responses that used the term light-year instead of years, showing that students do not understand that the term light-year is a measurement of distance and not a measurement of time. Other Partials were given to students who divided by two instead of multiplying by two in calculating the travel time. Although there were a range of Wrong responses (anyone who answered ‘yes’ it is possible to make a trip to the star at half light-speed in three weeks), explanations (when provided) included nonsensical calculations, appeals to advanced technology, and saying that it is possible in movies. The following student quote from the laboratory pretest is an example of confusion between the terms light-year and year and perhaps the speed of light as well:

> *“I would be able to make my trip to Sirius in less than 3 weeks, because I’m moving in ly [light-years] in which are extremely faster than standard years.”*

On the second part of the Laboratory 1 pretest, students were questioned about which objects would be reachable, from Earth, in three weeks if traveling at half of the speed of light (Tables 3.2, A3.2, $N = 74$). Nearly half of students (46%) answered correctly with object(s) in our Solar System. About a quarter (27%) received a Partial because they additionally listed objects such as stars and galaxies. Very few (5%) students were Correct in their explanation. Two-thirds did not provide an explanation.

Several weeks into the course, the term light-year and the mathematical relationship between distance, speed, and time continued to generate confusion for students, as evidenced by midterm exams. Similar to the pretest question, students were asked on Exam 1 to determine how long it would take to travel to distant objects at half of the speed of light (Tables 3.3, A3.3, $N = 65$). In this case, the students were not required to provide an explanation for their responses. Student responses coded as Correct rose to 40%, while Incomplete and Partials dropped to 5% and 6%, respectively. However, a substantial fraction (43%) of the responses was coded as Wrong. Common mistakes students continued to make included failing to double the distance in light-years to calculate travel time in years (but instead dividing by two), as well as using the term light-year as a measurement of time.

More students, about two-thirds (63%–66%), were Correct on Exam 1 when asked about traveling *at* the speed of light. We used a FIB question on travel time (Tables 3.4a, A3.4a, $N = 38$) as well as one on lookback time (Tables 3.4b, A3.4b, $N = 27$) to assess this concept.

We analyzed MC questions administered on the exams to further explore student ideas regarding the definition of the term light-year and its relationship to a year. On one question given on Exam 1 (Tables 3.5a, A3.5a, $N = 36$)

Table 3.3. Exam 1: Half speed of light (FIB).

N	C	I	P	W	T	NS	NR
65	**40**%	5%	6%	43%	0%	0%	4%

Table 3.4. Exam 1: Speed of light (FIB).

Question	N	C	I	P	W	T	NS	NR
a. Travel time	38	**66**%	5%	5%	21%	0%	0%	3%
b. Lookback time	27	**63**%	7%	7%	19%	0%	0%	4%

Table 3.5. Exam 1: Definition of light-year (MC).

Question	*N*	C	W	NR
a. v1	36	**97**%	3%	0%
b. v2	29	**72**%	25%	3%

about an advertisement not making sense because "light-years [are used to] talk about time, but a light-year is a unit of distance," nearly all of the students (97%) answered correctly. On another question on Exam 1, students were asked about viewing an object 1000 light-years away ($N = 29$). In this case, only about three-quarters (72%) answered correctly, giving a lookback time in years, while a fraction (10%) incorrectly chose a response that used the term light-year as a measurement of time (see Tables 3.5b, A3.5b).

When we probe students' ideas at the end of the course by analyzing a Final Exam question using the half speed of light (Tables 3.6, A3.6, $N = 48$), we see that only about a third of students correctly answer this question at the end of the semester, similar to the percentage after Exam 1 (40% Correct). The percentages of Incomplete responses were similar to those on Exam 1. The answers that were coded as Partial rose from 6% to 29% and those coded as Wrong dropped from 43% to 27% relative to Exam 1. Responses coded as Partial are indicative of confusion between the terms year and light-year. Responses coded as Wrong include mathematical errors relating to proportionality, other conceptual errors such as very short travel times, and may or may not include confusion between years and light-years. This implies that while the term light-year continued to be used by some students as a measurement of time at the end of the course, students' mathematical abilities improved and some other conceptual errors were cleared up.

For a FIB question on lookback time on the Final Exam (Tables 3.7, A3.7, $N = 39$), response rates for Correct and Incomplete were again similar to those on Exam 1, and again the number of responses coded as Partial rose (from 5%–7% to 23%) while the number of responses coded as Wrong decreased (from 19%–21% to 5%). Again, responses coded as Partial include confusion between the terms year and light-year. Responses coded as Wrong include conceptual errors such as very short travel times and may or may not include confusion between years and light-years, but generally do not include the mathematical errors relating to proportionality that plague responses to the half speed of light questions.

On the Final Exam ($N = 58$), students were administered the MC question relating distance in light-years and lookback time in years; 88% answered correctly and 3% incorrectly chose a response that included the term light-year (see Tables 3.8, A3.8).

Table 3.6. Final exam: Half speed of light (FIB).

N	C	I	P	W	T	NS	NR
48	**33**%	4%	29%	27%	0%	0%	3%

Table 3.7. Final exam: Speed of light lookback time (FIB).

N	C	I	P	W	T	NS	NR
39	**67**%	3%	23%	5%	0%	0%	3%

Table 3.8. Final exam: Definition of light-year (MC, v2).

N	C	W	NR
58	**88**%	12%	0%

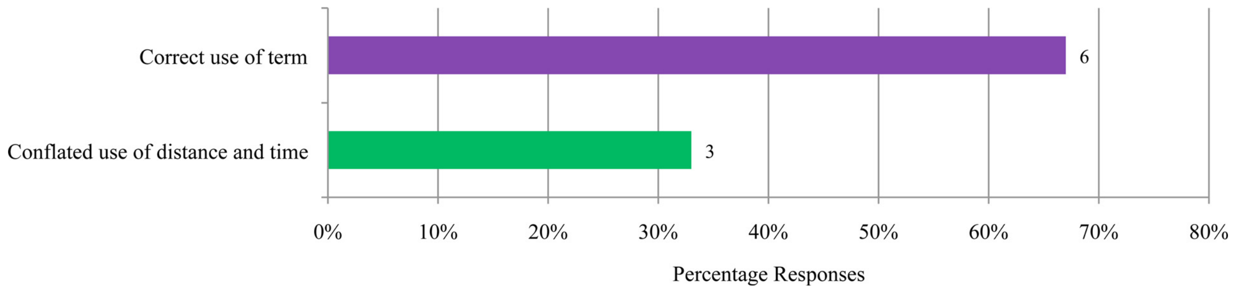


Figure 3.3. Interviews: Light-years. $N = 9$ students used the term light-year in interviews. Students were either asked specifically, or it came up during the course of the interview.

By thematically coding all speed of light and half speed of light FIB questions on Exam 1 ($N = 195$) and on the Final Exam ($N = 126$), we found that 11% of students incorrectly used the term light-year as a measurement of time on Exam 1, whereas 25% of students did so on the Final. This shows that more students confused the term light-year with the term "year" at the end of the semester than right after learning the material.

All interviews occurred after instruction on the definition of the light-year. Nine students were specifically asked about the term light-year or used the term during the course of their interviews (Figure 3.3). Consistent with exam data, two-thirds of the students correctly used the term light-year as a distance and one-third of the students used the term light-year as both a distance and a time.

Interview excerpts highlight some of the ways in which students may conflate the use of light-year as a distance and as a time. For example, a student ("O") correctly describes the relationship between light-years and travel time but incorrectly states that a light-year is a measure of distance and time:

> *Well, light-year is a measure of time, a distance or a measure of time, so something is 5 light-years away it would take 5 years to reach it, if it's 20 light-years it would take 20 years to reach it.*

When asked what the word light-year means, another student ("J") incorrectly describes it as a time:

> *Light-year means the time light takes to travel in a year.*

However, when prompted to elaborate, the student correctly describes the relationship between light-years and travel time:

> *[If] its ten light-years away that means that light has traveled ten years to get to where I'm at.*

3.1.2 Implications

We examined our data to catalog student ideas throughout the course. From the precourse homework essays, we see that respondents who addressed the term light-year were split between using it as a distance and as a time. Students were not required to discuss the term light-year, so these numbers are small (9 of 55 students, or 16%, addressed the topic). In the larger nationwide precourse survey, Bailey *et al.* (2012) found that 55% of students used the term light-year as a measure of distance, 28% identify it as a measure of time, and 9% as a speed. In addition, since responses in the Bailey *et al.* (2012) surveys can be coded as having more than one idea, about an eighth (12.4%) of the responses were identified as having an apparent contradiction—in other words, respondents seemed to hold more than one of these ideas (i.e., light-year as a measure of distance, time, or speed) at the same time. Results from our precourse homework essays and laboratory pretests also suggest that few students have a robust (i.e., full, clear) understanding of distances and travel times prior to instruction.

By comparing laboratory pretest, midterm, and final exam data, we see that learning gains are hard-fought and that students tend to revert back to the use of term light-year as a measurement of time by the end of the semester. Studies of student ideas in electricity and magnetism, tracked on finer timescales throughout the semester than in our data, show a similar reversion to preinstruction alternate conceptions (e.g., Sayre and Heckler 2009). In our case, it is possible that the reversion occurs because the term light-year being a distance measurement was stressed only during the first half of the semester. The results of one MC question seem to contradict this trend, with a slightly greater fraction of students answering correctly on the Final Exam than on Exam 1. However, by comparing types of exam questions, we see that, in general, students are more likely to use the term light-year correctly as a distance on MC questions than on FIB questions. Interviews confirm that student ideas are conflated as to whether a light-year is a distance or a time and how it relates to the speed of light. While two-thirds of interviewees used the term light-year correctly, one-third conflated distance and time. We speculate this might be because students have an easier time recognizing a selection from choices than recalling it themselves. This suggests that students' recognition of the unit light-year as a distance is superficial and that the concept should be continually reinforced explicitly (not just implicitly) throughout the course in order for students to achieve and maintain mastery of this concept.

Finally, students also seemed to be hindered by a difficulty with proportionality and the relationship between distance, speed, and time. Students fared worse on half speed of light questions than on speed of light questions. On the half speed of light questions, students often divided by two instead of multiplying by two in order to get the time given the distance. This could be because students are not triggering their physical intuition or might be thinking, "half is half." We suggest that instructors use conceptual analogies with everyday life, e.g., the fact that walking slower means it will take you longer (more time) to arrive at your destination.

3.2 Absolute and Relative Sizes and Distances of Objects

3.2.1 *Results*

We probed students' preinstructional ideas about relative and absolute distances with the precourse homework essays. Of the 55 students who completed the precourse homework essay, 53 students commented about the sizes or distances within the Universe (Figure 3.4). Many of these responses were nonspecific, relative terms. The most common responses were: that the Universe is infinite (36%) or vast/big (20%) or both infinite and vast (15%). Students gave a variety of other size and distance descriptors, such as "Earth is miniscule" (9%), objects are far apart (4%), the Sun is the biggest star, Jupiter is the largest planet, the Galaxy is "astronomical" in size, and more. Only three students put numbers to a description, one (incorrectly) saying that the Galaxy is 30,000 light-years across, one (approximately correctly) stating that the distance to the Andromeda Galaxy is 2 million light-years, and one (correctly) stating that the distance to the Sun is 93 million miles.

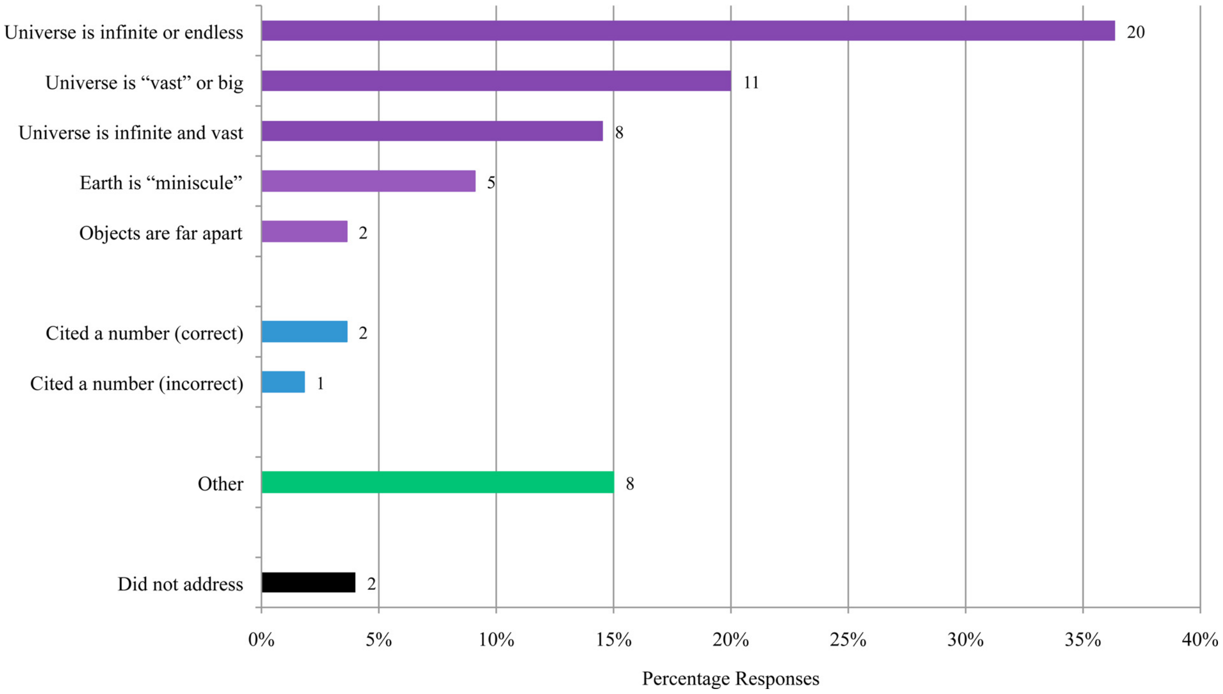


Figure 3.4. Precourse homework essay: sizes and distances. $N = 55$ essays were collected and analyzed for themes. It is possible for an essay to be coded with more than one theme. Since essays were free-response, the number shown for each category reflects a lower limit on the number of students who hold a particular concept. Many students described sizes and distances in their homework essays, but responses were often nonspecific relative terms.

In order to get a sense of students' ability to recognize absolute scales after the initial unit on size and scale, students were asked to match the distance or size of five objects with a choice of eight possible answers on Exam 1 (Tables 3.9, A3.9 and Figure 3.5). First, the question was coded with a point system, 0-5. One point was given for each correct object-size/distance match. Only about a third of students got at least 4 of the 5 correct. Next, the question was coded for thematic response frequency. The results show that the correct answers were the most popular for each object, but there was considerable spread.

In order to get a sense of students' ideas regarding relative sizes and distances after the initial unit on size and scale, we asked a number of ranking task questions on Exam 1. A distance-ranking question on Exam 1 included five objects: the Andromeda Galaxy, the Pleiades star cluster, the Hubble Space Telescope, Jupiter, and the Sun (Tables 3.10, A3.10, $N = 36$). This question was similarly coded with a 0-5 point system. One point was given for each object placed in the correct order. Slightly more than half of the students received a score of "5" for

Table 3.9. Exam 1: Matching absolute scales.

N	5	4	3	2	1	0	NR
65	**22%**	12%	15%	22%	20%	8%	2%

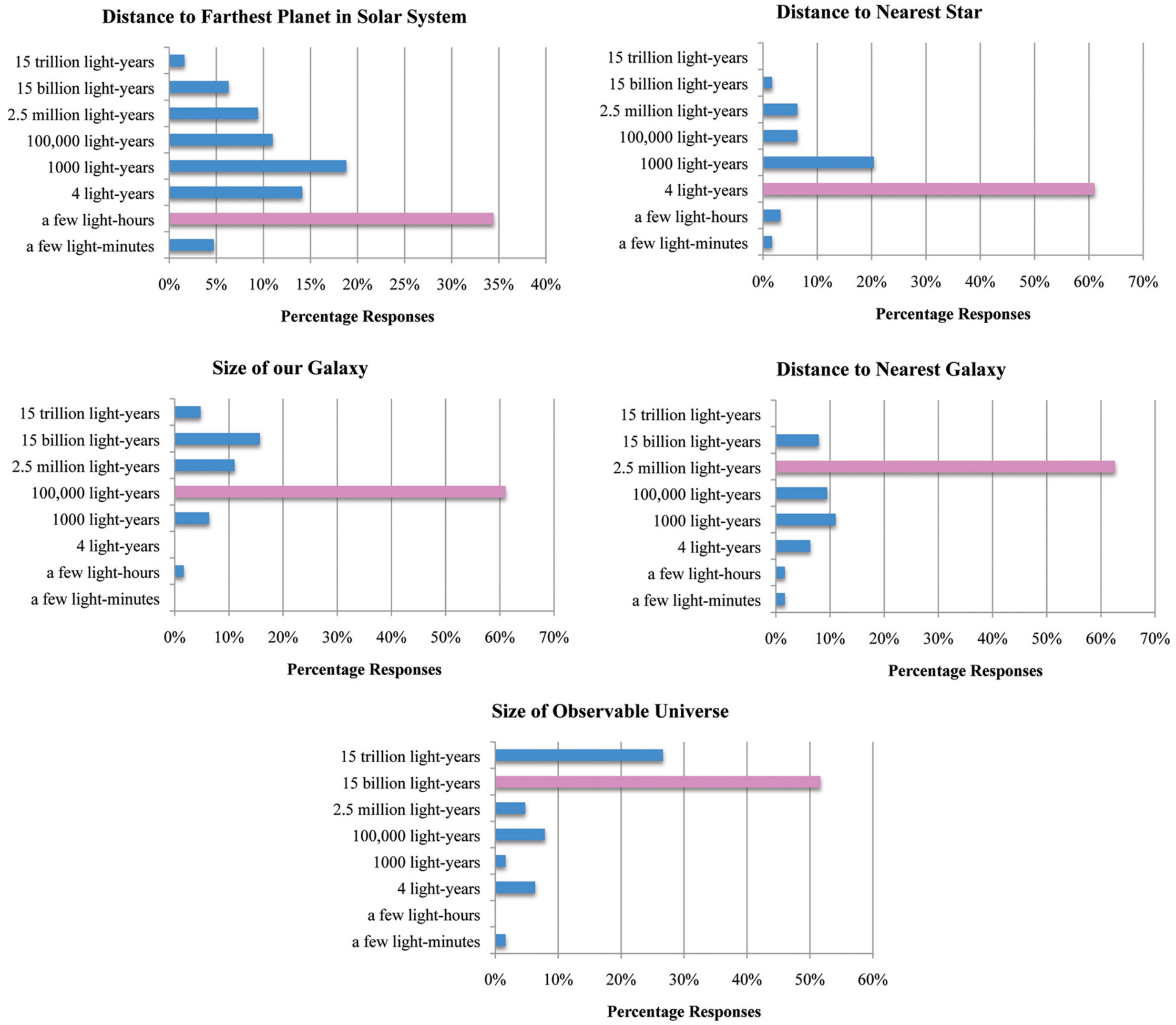


Figure 3.5. Exam 1: Matching absolute scales, thematic coding. The correct answer is highlighted in pink. Here, $N = 64$ because one student did not respond to the question.

Table 3.10. Exam 1: Object distance ranking task.

N	5	4	3	2	1	0	NR
36	**56%**	N/A[a]	36%	3%	6%	0%	0%

[a] Since this was a ranking task with 5 objects, it was not possible to score a 4 on this task.

placing all objects in the correct order and about a third received a score of "3" for ranking three objects correctly (note that a score of "4" is not possible here, since the task was to rank the distance—if one object was ranked incorrectly, at least one other object was then placed in the wrong order, too). Among the 16 students who scored 3 points or fewer, 10 students incorrectly ranked the star cluster to be farther than the Andromeda Galaxy and 6 students incorrectly ranked the Sun to be farther than Jupiter.

An open-ended size-ranking question on Exam 1 included seven objects (Tables 3.11, A3.11, $N = 29$). Rankings were coded with a 0-7 point system. In this case, a little over half (55%) of the students received the maximum score of "7". Explanations were coded for correctness, based on the nature of the objects involved (e.g., stars are bigger than planets, star clusters are bigger than single stars) and not exact numerical values for specific objects. About a quarter (24%) gave Correct explanations, while 31% were Incomplete, 21% Partial, and 14% Wrong.

In order to get a sense of students' ideas regarding relative sizes and distances at the end of the semester, we asked comparable questions on the Final Exam. The distance-ranking question (Tables 3.12, A3.12, $N = 32$) had a similar outcome in five-point responses on the Final as it did on Exam 1. However, out of 13 students who scored 3 points or less, the number of students who incorrectly ranked the star cluster to be farther than the Andromeda Galaxy dropped to 3, while the number who incorrectly ranked the Sun to be farther than Jupiter rose to 10.

A size-ranking question on the Final Exam (Tables 3.13, A3.13, $N = 26$) asked students to rank a subset of five of the original seven objects. About two-thirds of the students scored the maximum of "5" on a scale of 0-5. Among the 9 students who scored 3 points or less, 7 students ranked the Sun as smaller in size than Jupiter.

All interviews were conducted after the initial unit on size and distance scales. Fourteen students were questioned directly about size and distance scales. In coding for themes (Figure 3.6), we see that almost all discuss some sort of relative distance (93%), a little over half (57%) discuss an absolute number, and about a third (36%) discuss a scale model. One student did not respond.

Table 3.11. Exam 1: Object size ranking task.

a. Content

N	7	6	5	4	3	2	1	0	NR
29	**55%**	3%[a]	17%	14%	3%	3%	3%	0%	0%

b. Reasoning

N	C	I	P	W	T	NS	NR
29	**24%**	31%	21%	14%	10%	0%	0%

[a] Since this was a ranking task with 7 objects, it was not possible to score a 6 on this task, except in the case that a student omitted an object.

Table 3.12. Final exam: Object distance ranking task.

N	5	4	3	2	1	0	NR
32	**59%**	N/A[a]	22%	6%	13%	0%	0%

[a] Since this was a ranking task with 5 objects, it was not possible to score a 4 on this task.

Table 3.13. Final exam: Object size ranking task.

N	5	4	3	2	1	0	NR
26	**65%**	N/A[a]	19%	8%	8%	0%	0%

[a] Since this was a ranking task with 5 objects, it was not possible to score a 4 on this task.

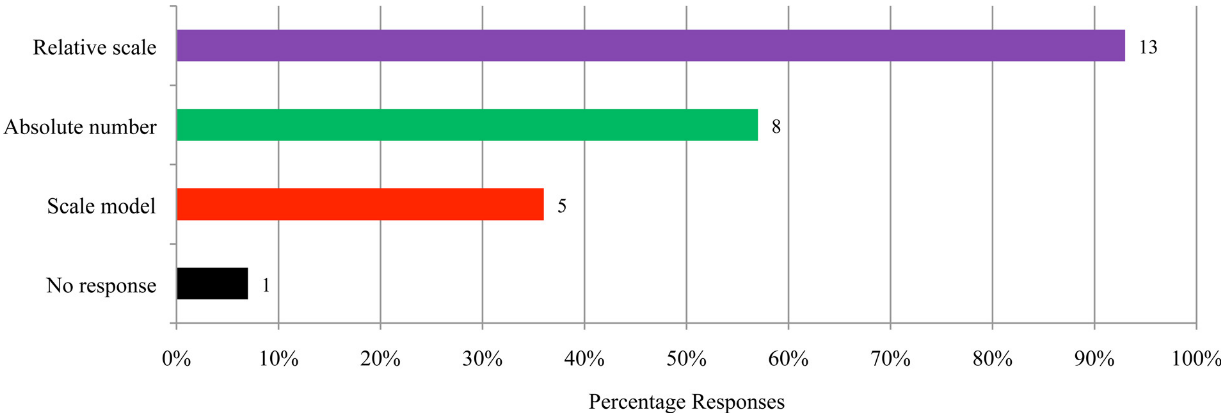


Figure 3.6. Interviews: $N = 14$ students were asked to discuss size and distance scales. Since responses were free-form, thematic codes can add up to more than 100%. All interviews were conducted after the initial unit on size and distance scales.

3.2.2 *Implications*

Exciting discoveries in astronomy (and other sciences) involve scales far removed from those of everyday human experiences, making the sizes and distances involved difficult for students to grasp. Students' ability to understand these discoveries will in turn depend on their ability to successfully navigate many orders of magnitude. Previous studies have examined students' ability to judge sizes and distances (prior to instruction). Tretter, Jones, and Minogue (2006) find that people are more accurate in judging sizes close to those on a human scale (us, things we use and see every day). The farther we go in either direction (smaller or larger), the worse people's accuracy. Miller and Brewer (2010) found a similar trend: undergraduates tend to overestimate the distance between Earth and the Moon, somewhat underestimate the distance between Earth and the Sun, and dramatically underestimate larger distances, such as those to the nearest star and the nearest galaxy. Insights from geoscience can shine light on students' inability to grasp large numbers. Cheek (2012) found that, "Fewer than half of the students [in this study] possessed knowledge of large numbers that was robust enough to enable them to understand processes in geologic time" (p. 1). Difficulties with orders of magnitude (e.g., thousands versus millions), proportional relationships, and problem solving strategies more generally can further contribute to the problem areas observed in Cheek's study as well as the present one. Delgado *et al.* (2007) stress the importance of having students establish connections among the various conceptions of size: ranking (ordering), grouping, ratio comparisons, and absolute scales and suggest that these conceptions might form a learning progression.

In examining students' descriptions of sizes and distances on the precourse homework essays, we see that while nearly all students (53 of 55, or 96%) addressed size and distance in some way, most used nonspecific relative words like "infinite," "vast," or "miniscule"; few students ($N = 3$, or 5%) cited absolute numbers with units. This supports the idea of relative scales as the start of a learning progression and highlights students' initial unfamiliarity with absolute scales.

After the introductory unit on size and scale, CSU students were given a question on the first midterm exam that assessed their ability to match absolute sizes or distances to five objects: the distance to the farthest planet in our Solar System, the distance to the nearest star, the size of our Galaxy, the distance to the nearest galaxy, and the size of the observable Universe. As seen in histograms (Figure 3.5), the correct answers were also the most popular answers when combined across semesters. However, there was considerable spread in the distributions, especially for the farthest planet in our Solar System. While this might seem at first glance to be contradictory to the work of Miller and Brewer (2010) because the farthest planet in our Solar System is the closest of the five objects listed, students might not yet have a robust concept of the Solar System, as we will see in Section 4 and as noted by Bailey *et al*. (2012). Some students could be overestimating the size of the Solar System by choosing distractors with larger-than-accurate values because “things in space are big.” Students could also be less familiar with the units of light-hours or light-minutes than light-years. Our data collection method also differed from that of Miller and Brewer (2010), who gathered open-response data and employed scale models when investigating this topic. We also see an interesting tail in the distribution for “size of the observable universe.” While the distribution peaks at the correct answer (15 billion light-years), a substantial number of students chose 15 trillion light-years. This effect might be related to the trends seen in the geoscience research on student difficulties distinguishing large numbers (Cheek 2012). Furthermore, when individual students were scored for correctness, only about a third of them scored a 4 or 5 on a 5-point scale for the task, suggesting that tasks with absolute scales are difficult for students.

Ranking tasks for size and distance on the first and final exams probed students’ ideas about the relative scales of objects. Between 55% and 65% of students were fully correct in their rankings. On the first exam, more students incorrectly ranked the distances of Galactic and extra-galactic objects, whereas on the final exam more students incorrectly ranked the distances of Solar System objects. There could be several explanations for this: Galactic and extra-galactic objects are revisited at the end of the semester, while Solar System objects are not, students could be confusing the sizes and distances of Solar System objects or students might believe that all of the planets are closer to each other than they are to the Sun. Furthermore, while the majority of students could rank objects correctly, many struggled to explain their size rankings on the first exam (only about a quarter were correct in their reasoning).

Results from our interviews support the findings from the precourse homework essays and exam responses that students have an easier time grasping relative scales. Specifically, in the interviews, students describe sizes and distances most frequently in relative terms, followed by citations of absolute scales and scale models. This lends support to the idea that relative scales are easier for students to grasp and might form the backbone of a learning progression.

3.3 Distance Measurement Techniques

Astronomers use what is called the “distance ladder,” in which a variety of techniques are used to measure the distances to objects increasingly farther away. In the minilecture used in our class at CSU, students are introduced to three of these: parallax, standard ruler techniques, and standard candle techniques (which use the inverse square law). The lecture includes a demo in which students view their thumb against a background of stars and observe that their thumb appears to shift more when it is closer to their eyes and less when it is farther away, illustrating the principle of parallax in everyday life. In lecture students also practice conceptual variations on the inverse square law (e.g., if a star of a given luminosity is three times farther away, it will be nine times dimmer, and variations thereof). In Laboratory 9: *Measuring Cosmic Distances,* students complete the *Lecture-Tutorial* on parallax where they measure the shift of a star from simulated star fields and compute the parallax; complete a variation on the University of Washington standard ruler laboratory (Palen, n.d.); complete a series of conceptual and numerical problems on the inverse square law; and have the option of measuring the distance to M100 using Cepheid variables (a variation on the ESA/ESO exercise; see Note-6). Again, verification-style laboratories were modified for use in the CSU course to allow students to also focus on making predictions and understanding concepts.

3.3.1 Results

In the precourse homework essays, students were not specifically prompted to describe how distances might be measured. However, nearly a quarter (24%) of students said that distances of or in the Universe could not be measured (Figure 3.7). For example, as one student describes,

> *I believe that the physical size and structure of the universe is so large that man cannot measure it in its entirety.*

---

[6] http://www.astroex.org/english/exercise2/.

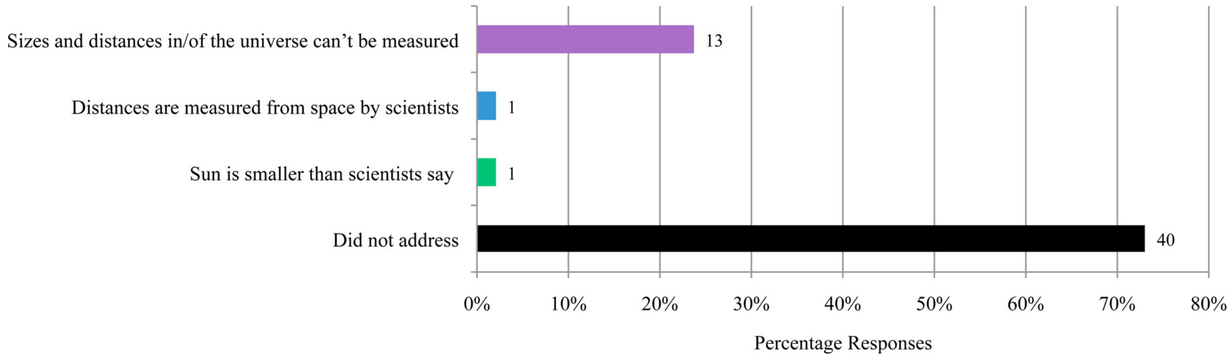


Figure 3.7. Precourse homework essay: size and distance measurement. $N = 55$ essays were collected and analyzed for themes. Since essays were free-response, the number shown for each category reflects a lower limit on the number of students who hold a particular concept.

Another student describes how they think scientists measure distances,

> *Scientists are discovering something everyday about the universe. They take trips to outer space and determined [sic] the distances from each other.*

One additional student describes their distrust of scientists' methods, explaining why they believe the Sun is smaller than scientists say,

> *I believe things in the universe are smaller than what scientists say. For example, I believe the Sun is smaller and closer than what is believed or taught by Scientists. I'm not claiming that my belief is the truth, as I'm rather guessing, but just that I believe the first principles in which scientists derived their facts are faulty.*

We are able to gauge student ideas about specific distance measurement techniques through other class artifacts. We first probe students' ideas on parallax following lecture but prior to laboratory through the pretest for Laboratory 9: *Measuring Cosmic Distances*. Only about a quarter (28%) of student responses were Correct on an open-ended question on parallax in the pretest (Figures 3.8 and 3.9, Tables 3.14, A3.14, $N = 36$). When citing their reasoning, students who answered incorrectly most often said that the star that was farther away shifted more. On Exam 3, a smaller but notable fraction of students still respond to a FIB version of the question incorrectly, with only slightly over half (51%) answering correctly (Tables 3.15a, A3.15a, $N = 56$). On Exam 3, students were also asked numerical reasoning MC questions about parallax (Tables 3.15b, A3.15b, $N = 35$; Tables 3.15c, A3.15c, $N = 21$). About half were Correct on these questions as well. On the Final Exam (Tables 3.16, A3.16, $N = 45$), 58% were Correct.

The laboratory pretest open-response question on the inverse square law yields some insight into students' preinstructional ideas on the standard candle aspect of the distance ladder (Tables 3.17a, A3.17a, $N = 36$). Only about a quarter (28%) of students' responses were Correct, whereas a substantial number of responses were Partial or Wrong. On Exam 3 (Tables 3.17b, A3.17b, $N = 56$) and the Final Exam (Tables 3.17c, A3.17c, $N = 32$), students

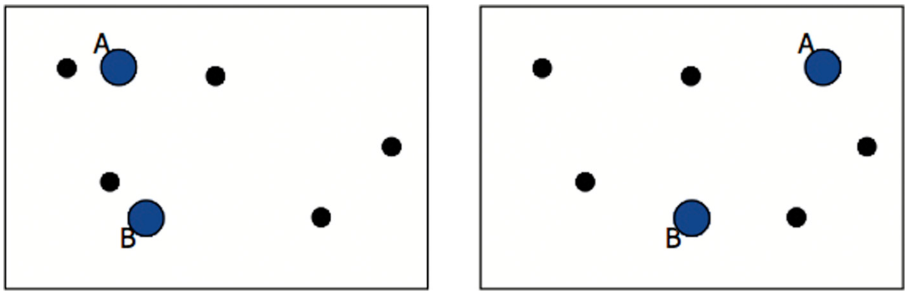


Figure 3.8. Figure used in the Laboratory 9 pretest question on parallax. Students were informed that the two pictures were taken six months apart and were asked which star is farther away, A or B? They were also asked to explain their reasoning. Similar illustrations were shown in FIB questions on the midterm and final exams (see Tables A3.15, A3.16).

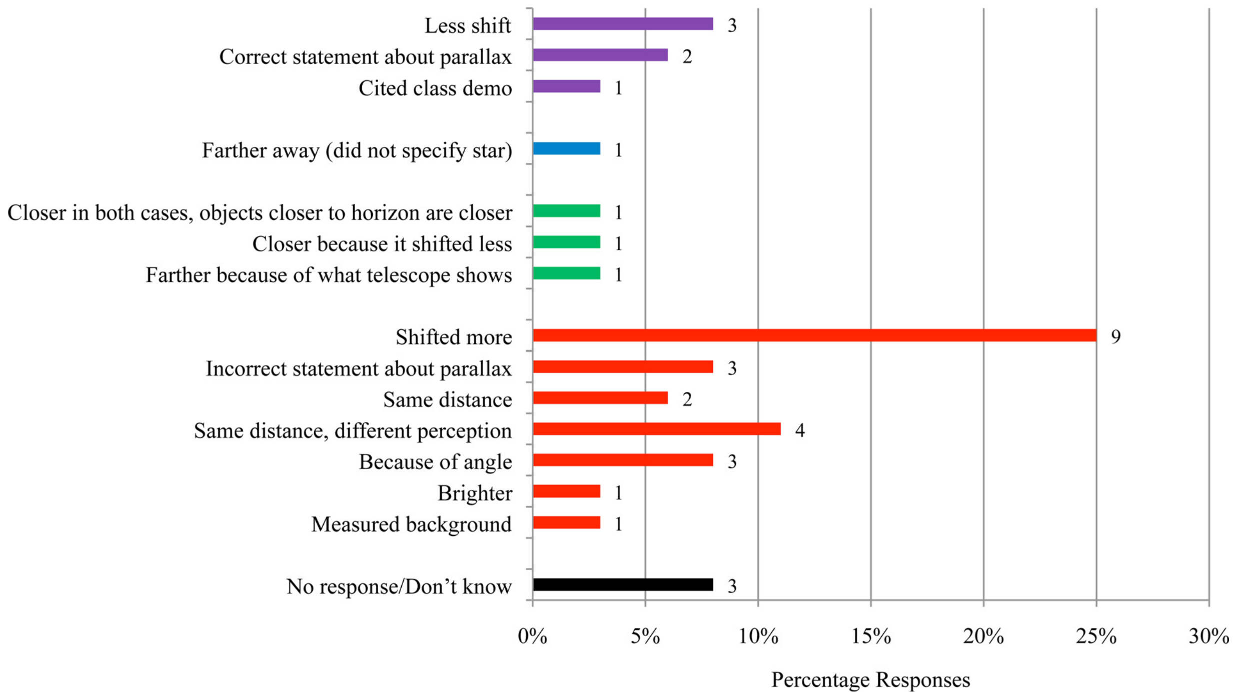


Figure 3.9. Laboratory 9 Pretest: Parallax measurements, reasoning, thematic frequencies. $N = 36$. Responses in purple are Correct, in blue are Incomplete, in green are Partial, and in red are Wrong. The correct answer and reasoning incorporated the elements: Star B is farther away because it shifted less.

Table 3.14. Laboratory 9 Pretest: Parallax shift.

N	A	B	SAME	NR
36	53%	**28%**	17%	3%

Table 3.15. Exam 3: Parallax.

Question	*N*	C	W	NR
a. Parallax shift	56	**51%**	45%	4%
b. Numerical (MC, v1)	35	**49%**	51%	0%
c. Numerical (MC, v2)	21	**57%**	43%	0%

Table 3.16. Final exam: Parallax shift.

N	C	W	NR
45	**58%**	40%	2%

Table 3.17. Inverse square law.

Instrument	*N*	C	I	P	W	T	NS	NR
a. Pretest	36	**28%**	14%	31%	22%	3%	0%	3%
b. Exam 3	56	**41%**	16%	21%	21%	0%	0%	3%
c. Final Exam	32	**59%**	16%	13%	13%	0%	0%	0%

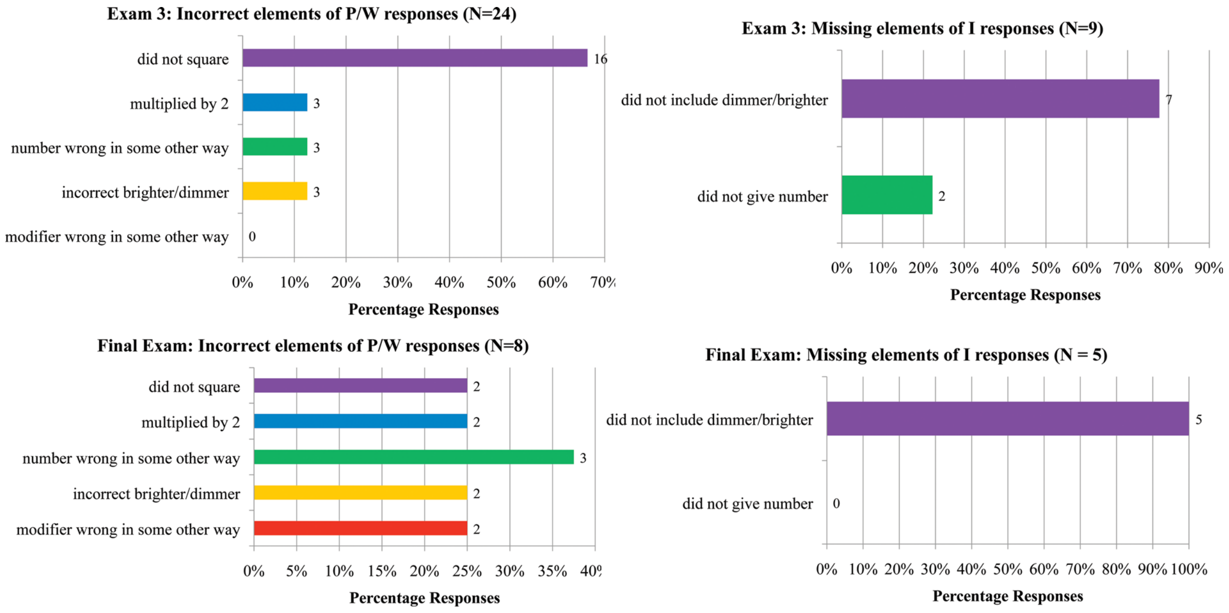


Figure 3.10. Exams: Difficulties with the inverse square law, thematic coding. It is possible for a response to be coded with more than one theme.

were asked FIB questions on the inverse square law (similar to the open-response question on the pretest). Students showed gains on both Exam 3, with 41% of responses Correct, and again on the Final Exam, with 59% Correct. The most common incorrect elements in the Partial and Wrong responses and the most common missing elements in the Incomplete responses are shown in Figure 3.10 for the exams. The most common mistake on Exam 3, by a substantial margin, was a mathematical error of not squaring the number given (67% of the responses classified as Partial or Wrong had this mistake). Other errors included multiplying by two, giving the wrong number in some other way, and using an incorrect brighter/dimmer modifier. On the Final Exam, the number of students who did not perform the squaring function was greatly reduced, becoming similar to the number of occurrences of other mistakes. Elements missing from Incomplete responses on Exam 3 were dominated by the lack of a brighter/dimmer modifier (78%), but also included not giving a number (22%). On the Final Exam, all Incomplete responses were coded as such because they did not include a brighter/dimmer modifier.

Parallax and the inverse square law were not discussed in interviews, either by direct questioning or in students' responses to other questions.

3.3.2 *Implications*

In the precourse homework essays, nearly a quarter of students stated that sizes or distances in or of the Universe "cannot be measured," with a subset of those specifying that this is because the sizes/distances involved are too large. This suggests that students might not understand the nature of scientific evidence or how it is possible to answer questions about the nature of the Universe. This is also seen with questions related to evidence for and measurements of the age of the Universe and the Big Bang (Bailey *et al*., 2012).

When examining our parallax data, we see that when explaining their answers on the laboratory pretest, students who were wrong most frequently gave the reason that the star that was farther away shifted more. Though we cannot say definitively, students might have been using the phenomenological primitive of "more is more" and "less is less" (diSessa 1988).

Students made conceptual gains on the inverse square law from pretest, to the third midterm, through the Final Exam. The pretest question on the inverse square law was asked prior to the lecture on the topic but was worded in such a way that students could construct the correct answer from the information given. Thus, the pretest was more of a probe of students' baseline mathematical reasoning abilities necessary for the task. The most common mistake on exams was the mathematical error of not squaring the number given. However, from Exam 3 to the Final Exam, more students correctly squared the number given when applying the inverse square law. The process of squaring a number is practiced throughout the course, for example, with the inverse square law of gravity and in the expression for kinetic energy (with the square of the velocity). Proportions in general are something practiced numerous times in the course, and FIB questions of various types of proportionality are common on CSU astronomy exams. Reinforcement of these ideas and skills via connections to other topics could account for the increase in student abilities from Exam 3 to the Final Exam.

A smaller number of students had trouble with the conceptual aspect that the light is dimmer if an object is farther away. However, we did not see an improvement in the number of students who correctly labeled their response with brighter/dimmer from Exam 3 to the Final Exam. These results suggest that both conceptual and numerical reasoning should be stressed when students are learning to do proportional reasoning from an equation.

4. STRUCTURE

In this section, we discuss students' understanding of structure in the Universe, focusing on the Solar System, Galaxy, Universe, and the relationships between them. For thousands of years, humans have been mapping the sky in an attempt to understand our place in the cosmos. We now understand the structure of the universe better than ever as a result of recent large optical surveys such as the Sloan Digital Sky Survey (Abazajian *et al*. 2009), which has mapped millions of objects, and 2dF (Cole *et al*. 2005), which has mapped hundreds of thousands. Surveys in nonvisible wavelengths further expand our knowledge of such structure (e.g., Fermi: Nolan *et al*. 2012; 2MASS: Skrutskie *et al*. 2006; ALFALFA: Haynes *et al*. 2011).

Understanding the structure of the Universe is one of the most fundamental themes students can take away from an astronomy class. Students should understand that our Solar System consists of not just the Sun and planets, but other, smaller, more numerous objects, such as comets and asteroids. They should also have a sense of the different types of planets and how that relates to how the Solar System formed. In moving beyond the Solar System, students should recognize that our Sun is just one of hundreds of billions of stars in our Galaxy; that is, galaxies contain solar systems. In addition to billions of stars, galaxies also contain gas, dust, and dark matter. There are three basic types of galaxies: spiral, elliptical, and irregular. Spiral galaxies feature a spherical bulge, a flattened disk with spiral arms, and a larger, spherical halo. Spiral galaxies typically have abundant gas and dust, ongoing star formation and a complement of younger, bluer stars. Elliptical galaxies are spherical to oblong in shape. They have less gas, dust and star formation than spiral galaxies and feature mostly older, redder stars. Galaxies themselves can be grouped into clusters and larger structures that make up the Universe. Understanding the hierarchical relationships between structures will allow students a framework for understanding new information about astronomical objects, both in class and beyond. At a minimum, we hope our students will understand that the Universe contains galaxies, which in turn contain solar systems.

As discussed in Section 3, during the first two weeks of class, students are introduced to objects in the Universe and how they are arranged through readings, lectures, ranking tasks, the laboratory on *Scales of the Universe*, and other activities. Additional activities later in the semester, such as *Toilet Paper Solar System* and the *Lecture-Tutorials* "Milky Way Scales" and "Galaxy Classification and Properties," help refine their understanding of these objects and the relationships between them. The structure of the Universe is a theme addressed throughout the semester.

4.1 Students' Preinstruction Ideas on Types of Objects in the Universe

Before looking at student ideas regarding individual structures in the Universe or the relationships between them, we begin by analyzing precourse homework essays ($N = 55$) for the types of objects students listed (Figure 4.1). This theme had a high response rate; 52 of 55 students (95%) listed astronomical objects in their essays. In our coding, we classified these into Solar System objects, Galactic objects, and extra-galactic objects, in addition to whether the students specifically used the words "stars," "galaxies," or "Milky Way." The closer the objects are to Earth, the more frequently students mentioned them. For example, 49 students (89%) mentioned Solar System objects, 16 students (29%) listed Galactic objects (other than stars) and 5 students (9%) mentioned extra-galactic objects. Again, we emphasize that these three categories are ours and not the students'. In their own words, 34 students (62%) mention "stars," 29 students (53%) mention "galaxies," and 13 students (24%) mention the "Milky Way" by name.

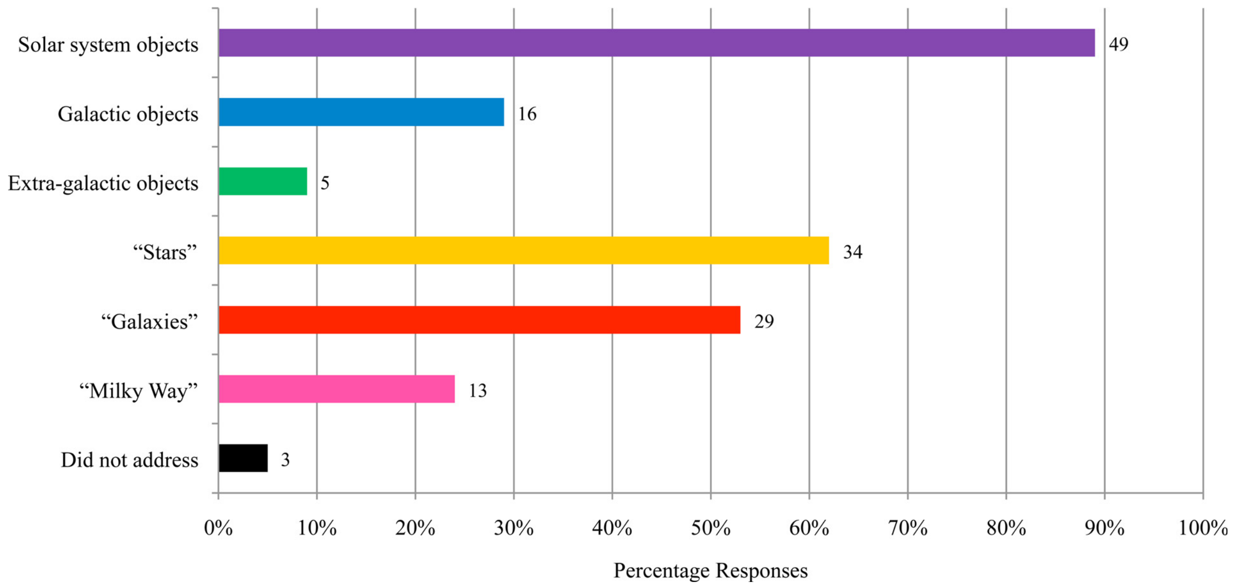


Figure 4.1. Precourse homework essay: Types of astronomical objects listed. $N = 55$ essays were collected and analyzed for themes. It is possible for an essay to be coded with more than one theme. Since essays were free-response, the number shown for each category reflects a lower limit on the number of students who hold a particular concept. We classified the objects that students listed into Solar System, Galactic (other than stars), and extra-galactic. Objects in quotes were ones that students specifically listed by name.

4.2 Solar System

4.2.1 Results

In Subsection 4.1, we see that students show a preference in their precourse homework essays toward listing Solar System objects. However, since the precourse homework essays are limited in what they can tell us about students' ideas of the structure of the Solar System, we decided to reanalyze the subset of precourse, open-response, written survey data collected by Bailey *et al.* (2012) from CSU students (Table 4.1, A4.1, $N = 17$). In the surveys, students were asked to describe the term "solar system." In order to make a comparison to later assessments (i.e., exam questions), we considered a response Correct if it included any three of the following components: Sun (or central star), planets, moons of planets, asteroids/comets. No students were fully correct, however, it should be noted that this precourse survey question was not worded in such a way as to require a minimum number of components. The vast majority of responses were Incomplete (71%); those were split between students who included one or two components. Fewer responses were Partial (12%) or Wrong (12%).

Students' postinstruction ideas about the structure of the Solar System were assessed with an open-ended essay question on the Final Exam (Tables 4.2, A4.2, $N = 48$). Students were asked to draw a diagram of the Solar System, label three major components, and to give a size scale. There were few fully Correct responses (17%),

Table 4.1. Precourse survey reanalysis: Solar System.

N	C	I	P	W	T	NS	NR
17	**0%**	71%	12%	12%	0%	0%	6%

Table 4.2. Final exam: Solar System structure (essay).

N	C	I	P	W	T	NS	NR
48	**17%**	38%	38%	6%	0%	0%	2%

but many were scored as Incomplete (33%) or Partial (42%). Only 6% were Wrong; these students drew the Galaxy instead of the Solar System. Among the incomplete responses, almost all (88%) were missing a size scale and more than half (56%) were missing one component. Of the responses that received a Partial, almost all (85%) had an error relating to size; about a third (35%) included a wrong component (e.g., noncentral stars, galaxies); and about a third drew incorrect diagrams (35%). When coding for themes (Figure 4.2), we see that overall nearly two-thirds (65%) of students had a missing, wrong or unclear size scale, nearly one-third (30%) were missing one component, several (17%) drew an unclear/incorrect diagram, and several (15%) included an incorrect component (noncentral stars, galaxies). Two examples of student responses are shown in Figure 4.3, including one typical incomplete response and one correct response.

We also used interviews to gain a deeper understanding of students' ideas about the Solar System. We asked the students to describe the term "solar system" ($N = 14$). All of these students were interviewed after the introductory material on structure and distance, as well as after in-depth instruction on the Solar System. The results are shown in Figure 4.4. All 14 students (100%) mentioned the Sun (or a central star) and planets.

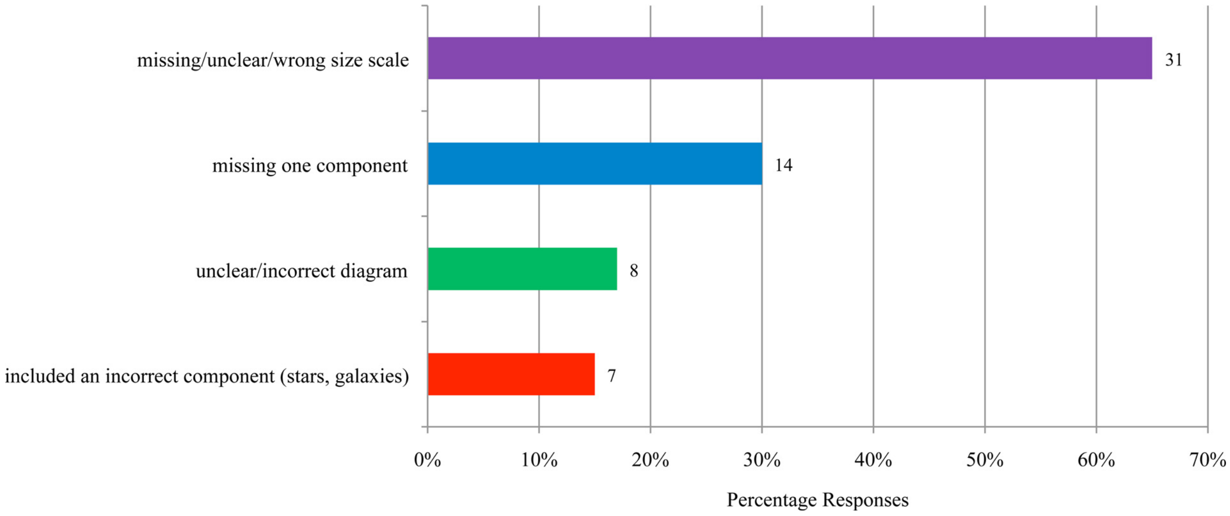


Figure 4.2. Final Exam: Difficulties with Solar System size and structure. $N = 48$ essays were analyzed for themes. It is possible for a response to be coded with more than one theme.

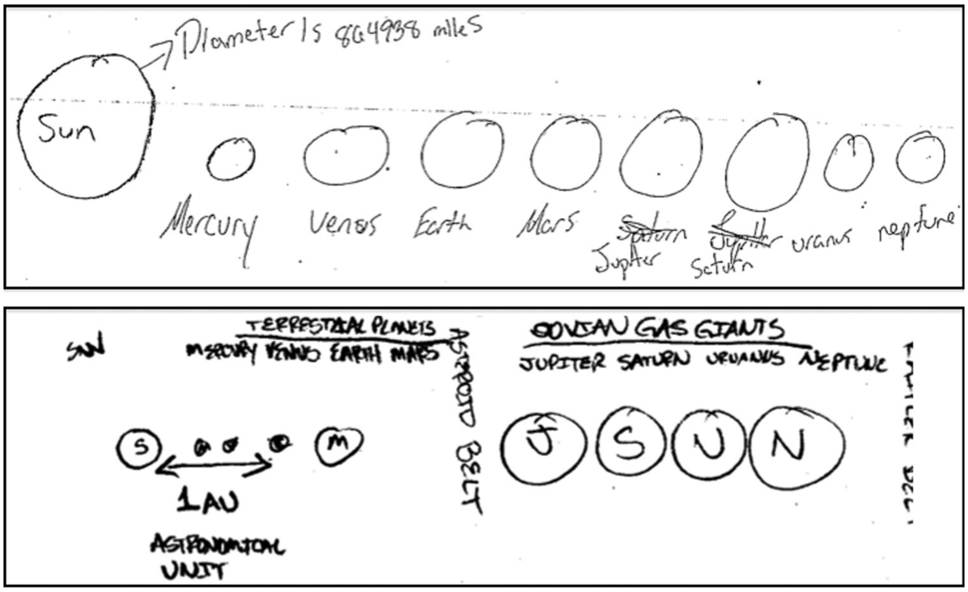


Figure 4.3. Examples of student drawings of the Solar System from the Final Exam. (a) This student drew only the Sun and named planets. (b) This student gives a much more detailed picture of the components in our Solar System.

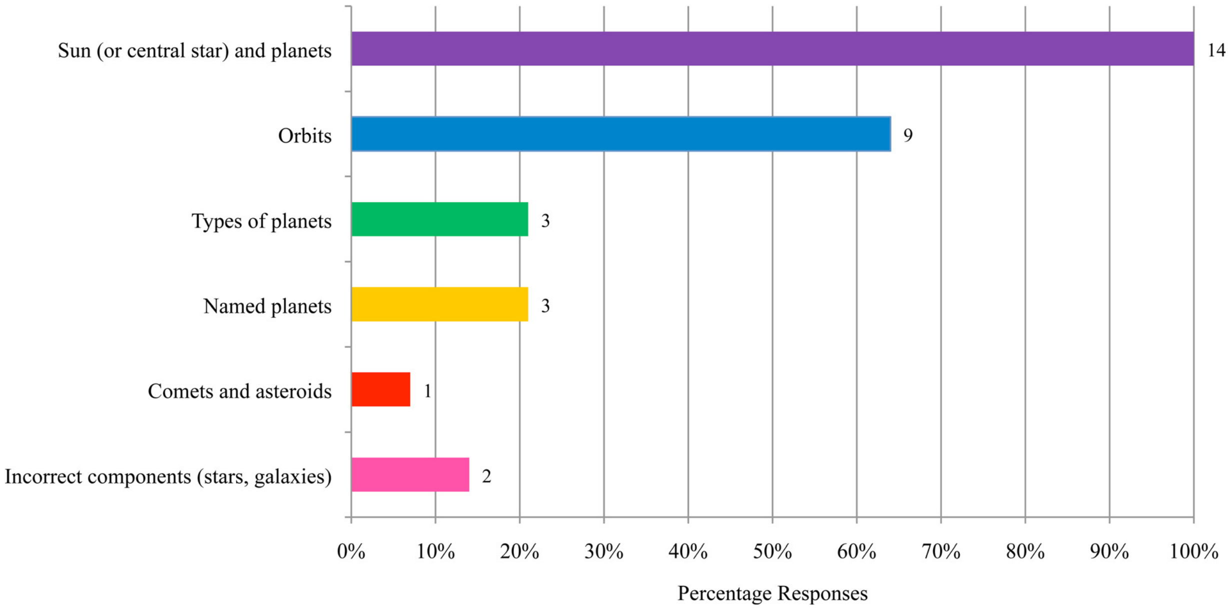


Figure 4.4. Interviews: Solar System. $N = 14$. Students were asked to describe the term "solar system." All of these students were interviewed after the introductory material on structure and distance as well as after in-depth instruction on the Solar System. The two students who listed incorrect components (noncentral stars, galaxies) had not yet been instructed in-depth on galaxies.

Furthermore, a majority (64%) mentioned orbits as an important feature (for example, planets orbiting the Sun). Of these, one student in an early interview gave an incorrect orbital relationship (planets orbit Earth), but later described the planets as orbiting the Sun. Other themes identified in students' responses were the planets by type (21%) and by name (21%). Only one student mentioned comets and asteroids when prompted if there was anything else he would like to add. Two of the early interviews (before in-depth instruction on galaxies) show ideas on structure that were not fully formed: one student thought there were galaxies inside the Solar System, and one student was debating whether there were other stars inside the Solar System.

Six of these students described that their previously held beliefs toward structure were biased toward only considering Solar System or other nearby objects. All three of the students who named the planets in detailed order recalled activities they did in elementary school on the planets. A fourth student said she previously thought that the Universe only consisted of Earth, the planets, and a few stars, based on what she had learned in elementary school. Another student said he only cared about the Solar System before taking the CSU course. A final student said he previously thought that the Galaxy was the biggest measure of everything. As one student ("N") explains:

> *"Like I kind of thought the galaxy was the Universe, and the Universe was something they just talked about like when kind of, you know constellation oh that's the Universe… I didn't know, I just thought it was referring to stars or something… I thought there was the Earth, the moon, and there was that sphere that encircled it. I don't know- my thoughts were very small about it."*

4.2.2 Implications

Student ideas about the structure of the Solar System are more robust (i.e., more well-formed, fuller, or clearer) at the end of the course than at the beginning. However, in both our exam data and interviews, we see that it is difficult for students to move beyond the idea of the Solar System as simply planets orbiting the Sun. When you consider some of the common astronomy activities at the K-12 level regarding the Solar System (e.g., Solar System travel brochures, book/Internet research of different planets), it is not surprising that students tend to focus on superficial factors such as names and order of the planets, rather than a more nuanced understanding of ideas in comparative planetology or about the smaller (but much more common) objects that orbit our Sun. Students in interviews cite the importance of their elementary education in shaping their mental models of the Solar System. We speculate that teaching students about the characteristics of Solar System objects within the physical context of what those characteristics tell us about how the Solar System formed might help place this information in a broader framework (e.g., Spencer and Guillaume 2006), so that students are not merely memorizing disjointed facts. Our current data do not allow us to investigate this possibility, however, it would be an interesting line of future research.

Furthermore, we see in exam questions that students are particularly challenged in assigning absolute size or distance scales to the Solar System, with only about one-third of students able to do so correctly. Relative scales (which were not asked on our exam question) are somewhat easier for students than absolute distance and size scales (Tretter, Jones, and Minogue 2006).

4.3 Galaxy

4.3.1 Results

In Subsection 4.1, we see that students show less of a preference in their precourse homework essays toward listing Galactic objects than Solar System objects. However, because the precourse homework essays are limited in what they can tell us about students' ideas of the structure of galaxies, we decided to reanalyze the subset of precourse, open-response, written survey data collected by Bailey *et al*. 2012 from CSU students (Table 4.3, A4.3, $N = 17$). In the surveys, students were asked to describe the term "galaxy." In order to make comparisons with exam assessments, we considered a response Correct if it included any three correct statements about the components of galaxies or that our Galaxy is the Milky Way. However, it should be noted that this precourse survey question was not worded in such a way as to require a minimum number of components. Only one response was Correct (6%). Nearly half (47%) of the responses were Incomplete and 18% were Partial. Only one response was Wrong (6%). However, nearly a quarter (24%) of students did not respond to this question.

Using exam data, we probed student ideas on the basic definition of a galaxy, the different types of galaxies, and the structure of spiral galaxies such as our own Milky Way. On Exam 3, we asked a series of MC questions on the number of stars in a galaxy (Tables 4.4, A4.4, $N = 26$) and galaxy types (Tables 4.5, A4.5, $N = 47$; Tables 4.6a, A4.6a, $N = 44$; Tables 4.6b, A4.6a, $N = 12$). We see that more than three-quarters (80%) of students correctly chose "billions" for the number of stars in a galaxy. Nearly all (91%) correctly answered a MC question about galaxy types, recognizing that "globular" is not a galaxy type. We also asked questions about the properties of galaxies where students could select among pictures of galaxies, choosing as many options as appropriate (as shown in Appendix Tables A4.6a and A4.6b). On the questions relating to the properties of spiral galaxies, slightly less than half of the responses (48%) were Correct, although few were Wrong (5%). For a similar question relating to the properties of elliptical galaxies, only about a third (33%) of students were Correct and a much larger fraction (42%) was Wrong.

Table 4.3. Precourse survey reanalysis: galaxy.

N	C	I	P	W	T	NS	NR
17	**6**%	47%	18%	6%	0%	0%	24%

Table 4.4. Exam 3: Number of stars in Milky Way (MC).

N	C	W	NR
26	**80**%	20%	0%

Table 4.5. Exam 3: Galaxy types (MC).

N	C	W	NR
47	**91**%	9%	0%

Table 4.6. Exam 3: Galaxy types.

Question	*N*	C	I	P	W	T	NR
a. Spirals	44	**48**%	34%	11%	5%	2%	0%
b. Ellipticals	12	**33**%	N/A[a]	25%	42%	0%	0%

[a] Due to the nature of the question asked, it was not possible to obtain an Incomplete on this item.

We also asked students two versions of a longer question related to Galactic structure on Exam 3. In one version of the question, we asked students to label a diagram of a galaxy (Tables 4.7a, A4.7a, $N = 26$). Only about a quarter of responses (23%) were Correct, but a large number were Incomplete (35%) or Partial (39%). Only a small fraction (4%) was Wrong. All of the Incomplete responses were coded as such because they were lacking a scale bar. All of the Partials had a missing, wrong, or unclear scale bar, and half had missing, wrong, or unclear labels for the components. When coding for themes overall, we see that about three-quarters (73%) of students had a missing, wrong, or unclear size scale and about one-quarter (23%) had a missing, wrong, or unclear halo component. In another version of the question (Tables 4.7b, A4.7b, $N = 21$), we asked students to sketch and label a picture of a galaxy, including the size scale and three components. Compared to the labeling-only version of this exam question, fewer responses were Correct (10%), with a substantial fraction again being Incomplete (14%) or Partial (57%). No students were Wrong. When coding for themes (in responses to this version of the exam question), we see again that problem areas included a missing, wrong, or unclear size scale (38%) and a missing, wrong, or unclear halo (52%).

We also asked a longer question related to Galactic structure on the Final Exam (Tables 4.8, A4.8, $N = 48$), similar to the essay question on Exam 3. Students were required to sketch the Galaxy, label three components and provide the approximate size of at least one component. In comparing results on Exam 3 and the Final Exam, we find that the fraction of Correct responses increased (to 29%), Incompletes increased (to 25%), Partials dropped (to 39%), and the number of Wrong responses remained small (4%). When coding for themes, we see the same problem areas again; overall half (50%) of the responses still had a missing, wrong, or unclear size scale and nearly half (46%) had a missing, wrong, or unclear halo. Only one student drew the Solar System instead of the Galaxy. Thematic coding of student difficulties on the midterm and final exams is shown in Figure 4.5. Two examples of student responses are shown in Figure 4.6, including one response with an unclear halo component and one correct response.

On the Final Exam, we also asked a MC question in which students had to choose the best description of a galaxy (Table 4.9, A4.9, $N = 58$). About three-quarters (78%) answered correctly.

All 15 students interviewed were asked to describe the term “galaxy.” All of these students were interviewed after the introductory material on structure and distance as well as after in-depth instruction on the Solar System. Furthermore, 13 had received in-depth instruction on the Milky Way Galaxy and 12 had received in-depth instruction on the different galaxy types and their properties. The results are shown in Figure 4.7. The student interview responses for “galaxy” were more varied than those for “solar system.” The most common description

Table 4.7. Exam 3: Galaxy structure.

Question type	*N*	C	I	P	W	T	NS	NR
a. Label (FIB)	26	**23**%	35%	39%	4%	0%	0%	0%
b. Sketch (Essay)	21	**10**%	14%	57%	0%	10%	0%	10%

Table 4.8. Final Exam: Galaxy structure, sketch and label (Essay).

N	C	I	P	W	T	NS	NR
48	**29**%	25%	39%	4%	0%	0%	4%

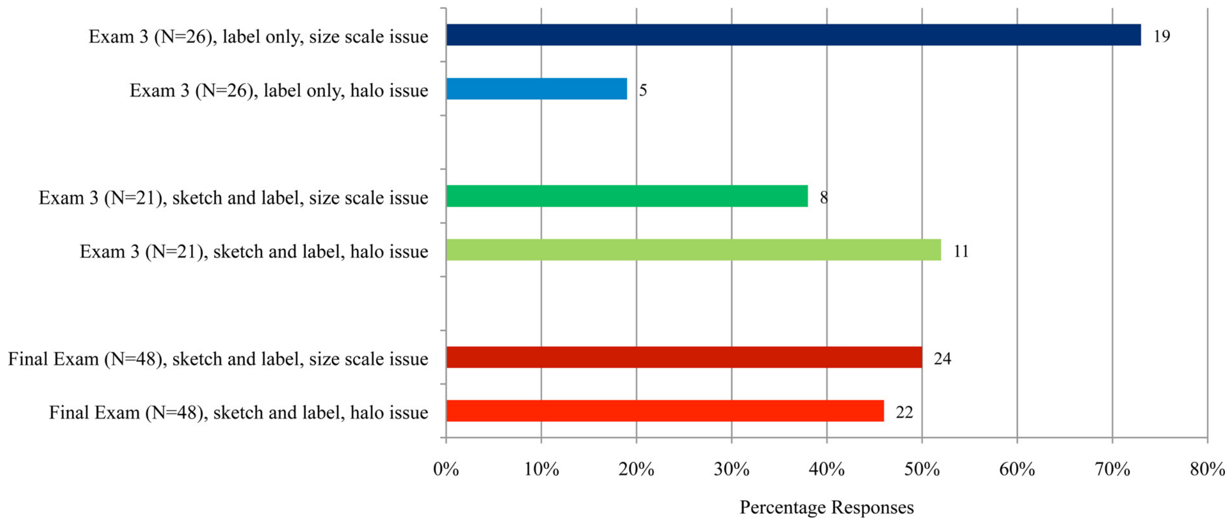


Figure 4.5. Exams: Difficulties with Galactic size and structure, thematic coding. It is possible for a response to be coded with more than one theme.

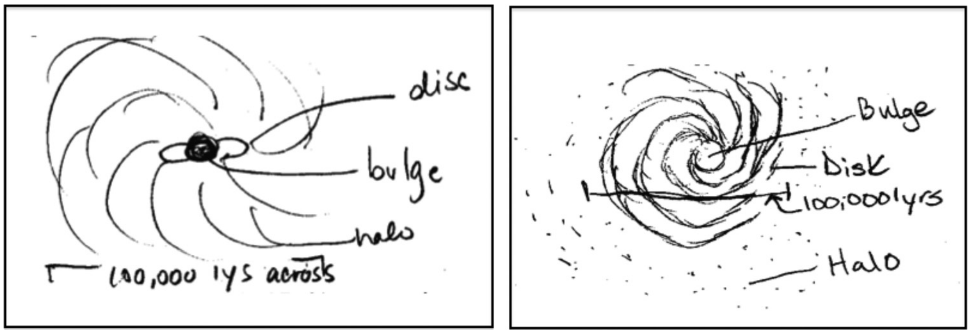


Figure 4.6. Examples of student drawings of the Galaxy from the Final Exam. (a) Like many, this student is unclear about the halo. (b) This student gives a correct picture of our Galaxy.

Table 4.9. Final exam: Definition of galaxy (MC).

N	C	W	NR
58	**78%**	21%	2%

of a galaxy was that it contains stars (73%). Of the 11 students describing the galaxy in that way, 4 specifically mentioned billions of stars, while 7 students gave descriptions such as “many,” “a lot,” or “trillions,” or simply said “stars.” The next most common theme was to list types of galaxies (40%), with fewer students (13%) going into detail of the properties of each type of galaxy. Four students (27%) named the Milky Way specifically. Four students (27%) described the bulge, disk, and halo of our Galaxy in varying degrees of detail. Only one student listed dark matter as a component of galaxies; this student was interviewed immediately after the laboratory on dark matter. In an early interview (after instruction on the Milky Way Galaxy, prior to instruction on galaxies), one student named planets in our Solar System; this student was clearly conflating the terms Solar System and Galaxy.

4.3.2 Implications

Students successfully answered short exam questions about the basic definition of a galaxy and the number of stars in a galaxy. These results are reinforced by our interviews. Students also easily recognized the names of the

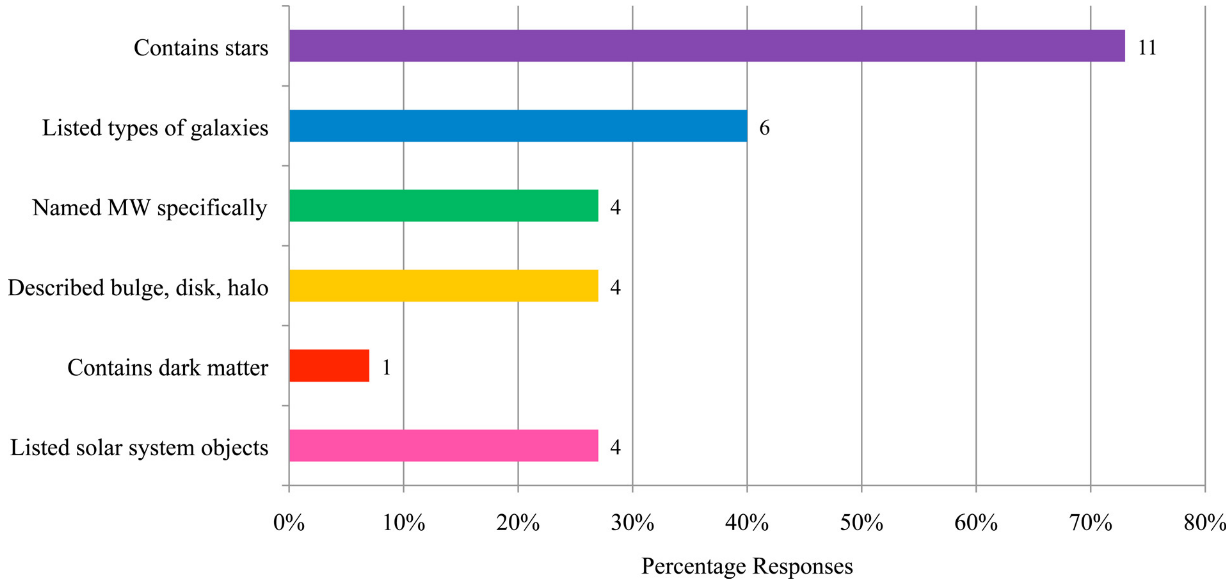


Figure 4.7. Interviews: Galaxy. $N = 15$. Students were asked to describe the term 'galaxy.' All of these students were interviewed after the introductory material on structure and distance as well as after in-depth instruction on the Solar System; 13 had received in-depth instruction on the Milky Way Galaxy and 12 had received in-depth instruction on the types and properties of galaxies in general.

three major types of galaxies on a short exam question. However, when viewing pictures of spiral and elliptical galaxies on exams, students had difficulty associating the picture with the properties. In interviews, 40% of students listed the types of galaxies, but fewer (13%) went into detail about their properties. This suggests that while students can fairly easily recall superficial definitions for a galaxy, they struggle with understanding the properties of different galaxy types.

In analyzing the CSU subset of precourse surveys, we see that students enter the course with a weak understanding of galactic structure. Student ideas about galactic structure are more robust on midterm exams than preinstruction (as would be expected). Furthermore, by comparing questions on Galactic structure on the midterm and final exams where students were required to provide their own drawing, we also see that student understanding continued to improve throughout the semester.

The biggest conceptual difficulties students seemed to face were the size of the Galaxy and both inclusion and properties of the galactic halo. On exams, only about a third to half of the students could provide correct scales (again consistent with student difficulties on absolute size and scale more generally; see e.g., Tretter, Jones, and Minogue 2006, Delgado *et al*. 2007). Only about half could correctly draw the halo, although about three-fourths could correctly label the halo if it was provided in a drawing. In interviews, only about a quarter of students chose to discuss the bulge, disk, and halo components of the galaxy, with varying degrees of success. While we simply prompted them to discuss the term galaxy and did not require them to discuss the components, the superficial level of detail provided by most students suggests that this is a difficult subject.

In interviews only one student (7%) mentioned dark matter as a component of galaxies; this student was interviewed immediately after a laboratory on dark matter in spiral galaxies. Similarly, only two students (4%) mention dark matter as a component of galaxies on the final exam (although they were not specifically prompted). This, combined with student difficulty visualizing the halo of a galaxy, suggests that instructors should consider bringing up the concept of the dark matter halo early in the course.

As with the Solar System, it appears that just labeling and naming the components of the Galaxy does not lead to a deep understanding of what they are. Teaching a more intricate story of what we observe and how that helps us understand the histories of galaxies could improve student understanding by anchoring names and labels to physical processes and putting them in a broader context (e.g., Spencer and Guillaume 2006). Our current data do not allow us to investigate this possibility, however, it would be an interesting line of future research.

4.4 Universe

4.4.1 Results

Several prominent ideas emerged when we examined how students described the Universe as a whole in the precourse homework essays (Figure 4.8). Students generally perceive the Universe as vast, empty, and dark. Specifically, the most common description fell into the category of having "empty spaces" that are vast, big, or large (33%). Students also describe the Universe as dark (18%), containing "everything" (11%), and being "indescribable" (7%). Fewer students mentioned that space has air/an atmosphere (5%), that there was *not* a lot of empty space (5%), or that it was quiet (2%). These responses are student descriptions of the Universe as a whole, and are in addition to any objects or "stuff" students might have listed as being in the Universe.

We also examined the precourse homework essays to determine students' ideas regarding the possibility of an edge to the Universe (Figure 4.9). Only three students (6%) describe the Universe as having an edge, whereas the majority who responded (38%) describe the Universe as "endless," "infinite," or specifically state the Universe has no edge. Of the eight students (15%) who gave some other response, three students (6%) made the distinction between an edge to "what we could see" and an edge to the entire Universe. One student described an end to galaxies but not empty space. Two students described a round Universe.

In interviews, we asked students to describe the term Universe (Figure 4.10, $N = 14$). In describing the Universe as a whole, students most commonly used the word "everything" (50%). The "vast" or "big" size of the Universe was popular, with 29% of students describing it as such; half of those students described it as "endless" or having "no boundaries" and one student described it as "larger than we can imagine." Again, these responses are student descriptions of the Universe as a whole, and are in addition to any objects or "stuff" students might have listed as being in the Universe.

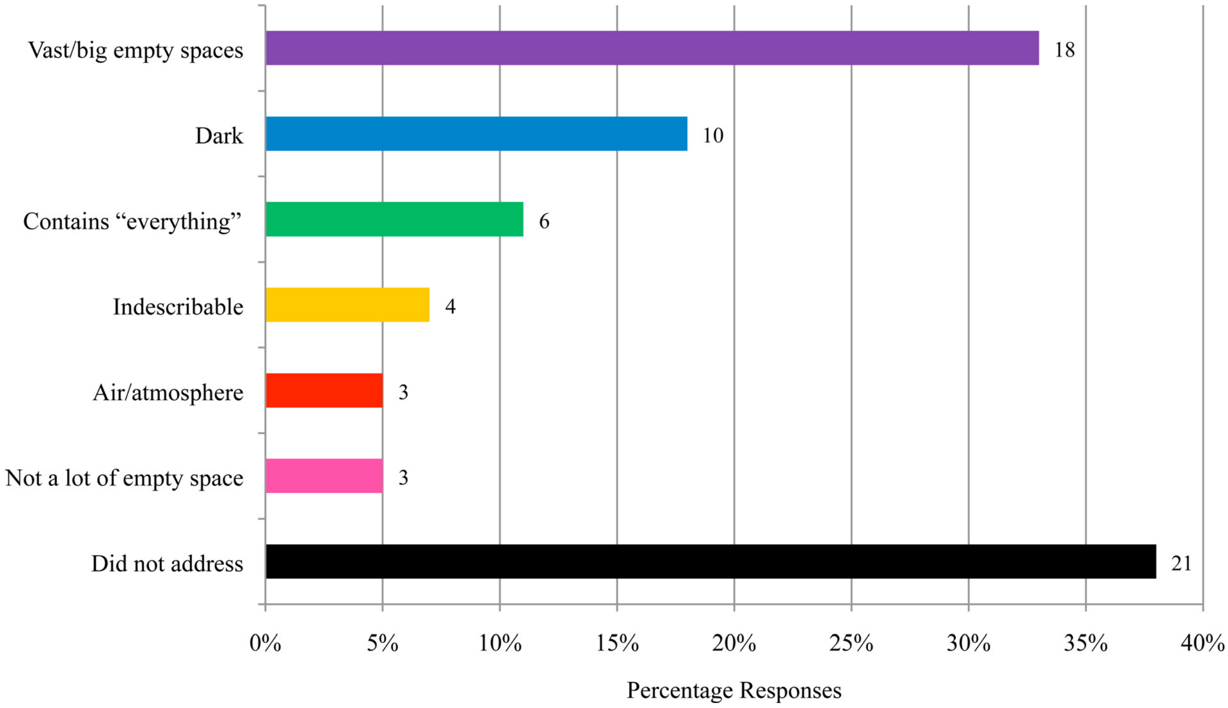


Figure 4.8. Precourse homework essay: Descriptions of the universe as a whole. $N = 55$ essays were collected and analyzed for themes. It is possible for an essay to be coded with more than one theme. Since essays were free-response, the number shown for each category reflects a lower limit on the number of students who hold a particular concept. These descriptions are of the universe as a whole and are in addition to any objects or "stuff" in the universe that the students might have listed.

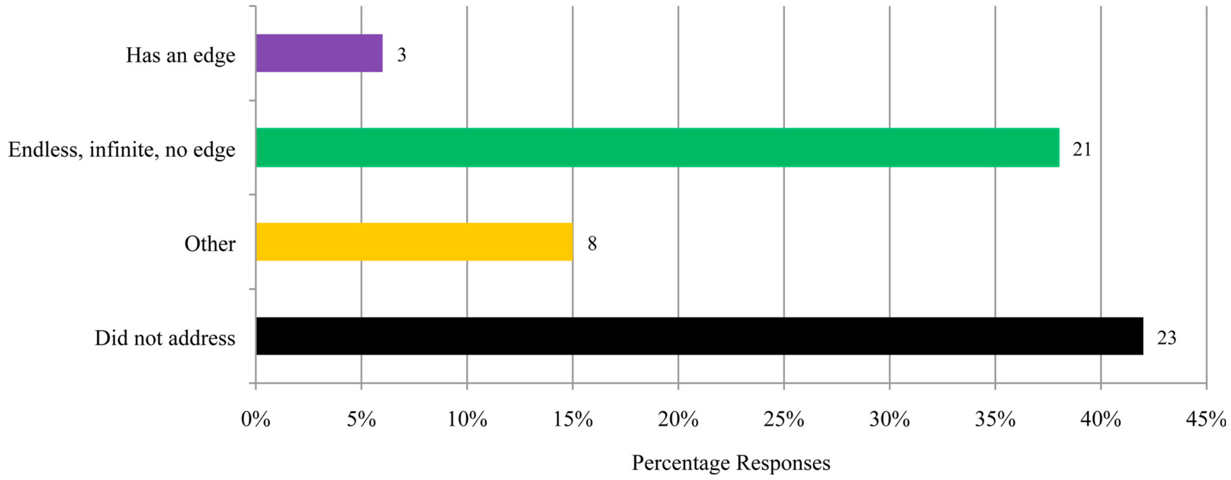


Figure 4.9. Precourse homework essay: Whether the universe has an edge or is endless. $N = 55$ essays were collected and analyzed for themes. It is possible for an essay to be coded with more than one theme. Since essays were free-response, the number shown for each category reflects a lower limit on the number of students who hold a particular concept.

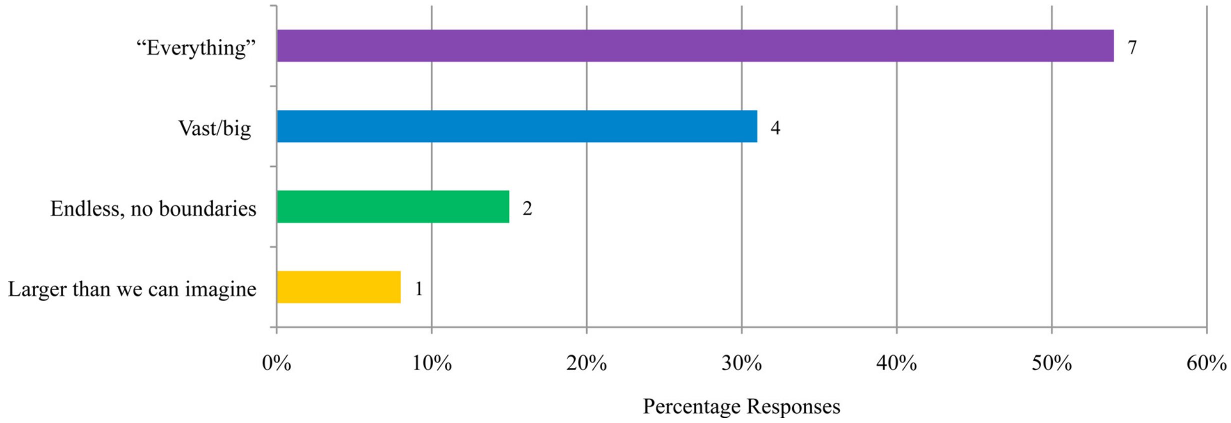


Figure 4.10. Interviews: Descriptions of the universe as a whole. $N = 13$. Students were asked to describe the term 'universe.' All of these students were interviewed after the introductory material on structure and distance as well as after in-depth instruction on the Solar System; 10 had received in-depth instruction on the types and properties of galaxies and 7 had received in-depth instruction on Big Bang cosmology. These descriptions are of the universe as a whole and are in addition to any objects or "stuff" in the universe that the students might have listed.

4.4.2 Implications

In our precourse homework essays, students seem to recognize that there are great regions of empty space in the Universe and that space is generally dark. A greater percentage of students in interviews describe the Universe as encompassing "everything," as we would hope as a course progresses. "Universe" is a term that gets used in everyday language (e.g., "universal remote") and so instructors might assume students enter the course knowing what it means. We often skip over actually defining it well, or at all, from a scientific standpoint (in Bailey *et al.* 2012, several different textbook definitions for the term "Universe" are listed, and one of the textbooks mentioned does not define the term).

When investigating students' ideas regarding an edge to the Universe, Wallace, Prather, and Duncan (2012) found that "many students reject the idea of an edge where the entire universe and all of existence come to an end; however, many (34–51%) believe the universe has an edge in the sense that the distribution of galaxies eventually ends, leaving only empty space." While we asked about the edge using somewhat different questions than this group in our precourse homework essays, we also find that few students (6%) think of the Universe as having a finite edge and that some students distinguished between an edge of all existence, an edge to what we could see, and an edge to matter.

4.5 Relationships Between Structures

4.5.1 Results

When we examined the precourse homework essays for how students thought objects in the Universe are arranged, we found that 44 out of 55 students (80%) addressed this topic (Figure 4.11). The most common student description (36%) was the concept of orbits, for example planets orbiting the Sun, moons orbiting planets, etc. A few students (11%) recognized that gravity or other physical mechanisms were the reason behind orbits and other phenomena. One student gave a vivid analogy for how structures organize due to gravity,

> *"These gaseous and solid bodies would then gather together, again thanks to gravity, to create small communities like towns and villages."*

A popular description of the Universe, given by 22% of students, is that is "chaotic," "random," or "swirly," or as one student put it,

> *"I would probably see many different masses of objects in various shapes, sizes, and colors that I would not be able to define. I can compare it to what I see when I look into a kaleidoscope that contains various shapes, sizes, and colors."*

This latter group of students did not have a robust scientific concept of astronomical structures at the start of the course.

Smaller numbers of students said there was a pattern to how objects are arranged but did not provide details (13%) or thought objects were spread out uniformly across the Universe (7%). There were a variety of other responses (27%), but each by very few students: from "every object has a purpose for its position" to the pattern is "explained by astronomers" to "someone thought about how they wanted everything to be in place."

We also examined the precourse homework essays to determine whether students have a sense of the hierarchical structure of the Universe preinstruction (Tables 4.10, A4.10, $N = 55$). For example, the Universe contains galaxies, which contain solar systems, which typically consist of a star and its planets and other debris. Only 21 of 55 students (38%) described some sort of hierarchical structure. When we categorized the student responses for correctness, we found that 22% of responses were Correct, few responses were Incomplete, Partial, or Wrong, and the majority of students did not respond. Examples of Incomplete responses included "planets are smaller

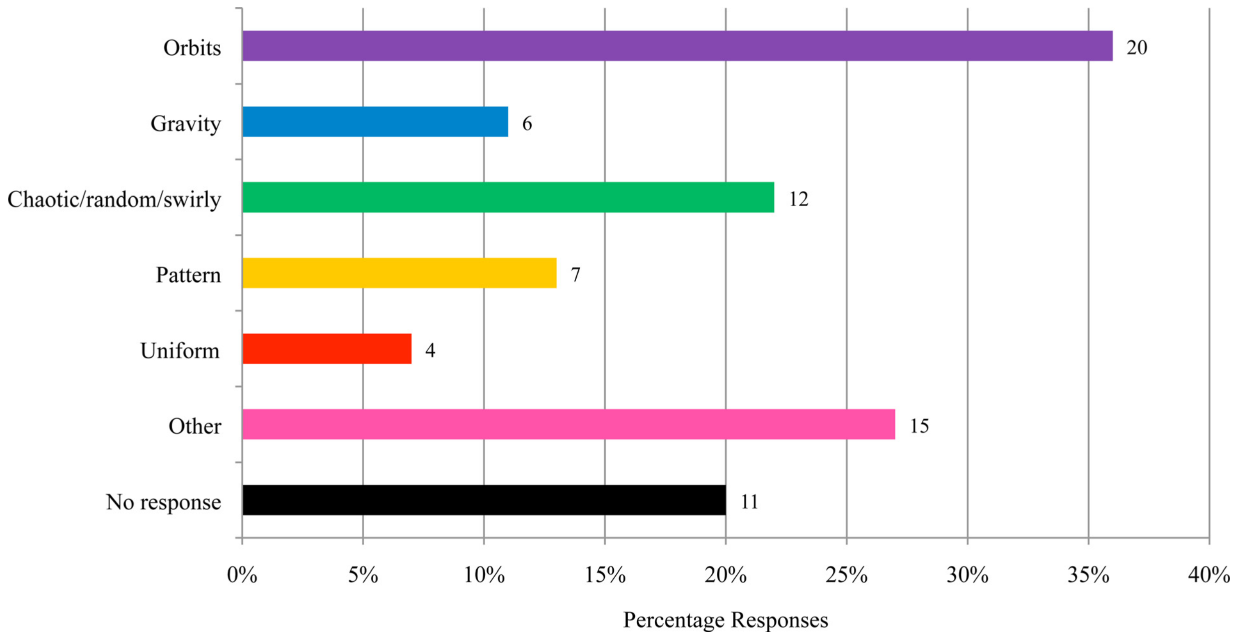


Figure 4.11. Pre-course homework essay: How objects are arranged. $N = 55$ essays were collected and analyzed for themes. It is possible for an essay to be coded with more than one theme. Since essays were free-response, the number shown for each category reflects a lower limit on the number of students who hold a particular concept.

Table 4.10. Precourse homework essay: Hierarchical structure.

N	C	I	P	W	NR
55	**22**%	4%	7%	6%	62%

than galaxies"; Partial responses included naming the planets in distance order and then explaining that "everything in our galaxy orbits the Sun"; and Wrong responses included "Earth is the center of the Universe."

We can compare these data to the subset of precourse, open-response, written survey data collected by Bailey *et al*. 2012 from CSU students (Table 4.11, A4.11, $N = 17$). In the surveys, students were asked to describe how the solar system, galaxy, and universe are related. In reanalyzing these data for the present study, we consider a response Correct if the students correctly described the universe as containing galaxies, which in turn contain solar systems. While the response rate was higher than that of the precourse homework essays, more than a quarter of students (29%) did not answer the question. Responses were fairly evenly split between Correct (18%), Incomplete (18%), and Partial (24%), with no Wrong responses. A few (12%) of responses were True but irrelevant.

To probe ideas at various points during the course, we gave students several ranking tasks related to hierarchical structure on exams. On Exam 1 (Tables 4.12a, A4.12a, $N = 53$), students were asked to rank by size: "Galaxy, Solar System, Universe, Earth" or "Galaxy, Solar System, Universe." A great majority (83%) were Correct in their ranking and very few were Wrong (2%). Responses coded as Partial (15%) meant students incorrectly ranked some but not all of the items. When asked to explain their reasoning in ranking Galaxy, Solar System, Universe (Tables 4.12b, A4.12b, $N = 36$), only half were Correct, although few (3%) were Wrong. A large number were Incomplete (31%), with responses that reiterated the ranking but did not describe each item or the relationships in detail. Responses coded as Partial (17%) included explanations that conflated two items, for example, "the Solar System is the Milky Way" or "the Milky Way Galaxy is only one Galaxy within our Solar System." On the Final Exam (Tables 4.13, A4.13, $N = 48$) when asked which was bigger, the Solar System or the Galaxy, over three-quarters of students (77%) were able to correctly choose Galaxy and explain their reasoning; only 2% were Wrong.

Table 4.11. Precourse survey reanalysis: Relationships.

N	C	I	P	W	T	NS	NR
17	**18**%	18%	24%	0%	12%	0%	29%

Table 4.12. Exam 1: Structure ranking task.

Question part	*N*	C	I	P	W	T	NS	NR
a. Content	53	**83**%	0%	15%	2%	0%	0%	0%
b. Reasoning	36	**50**%	31%	17%	3%	0%	0%	0%

Table 4.13. Final Exam: Solar system and galaxy ranking, including reasoning.

N	C	I	P	W	T	NS	NR
48	**77**%	4%	17%	2%	0%	0%	0%

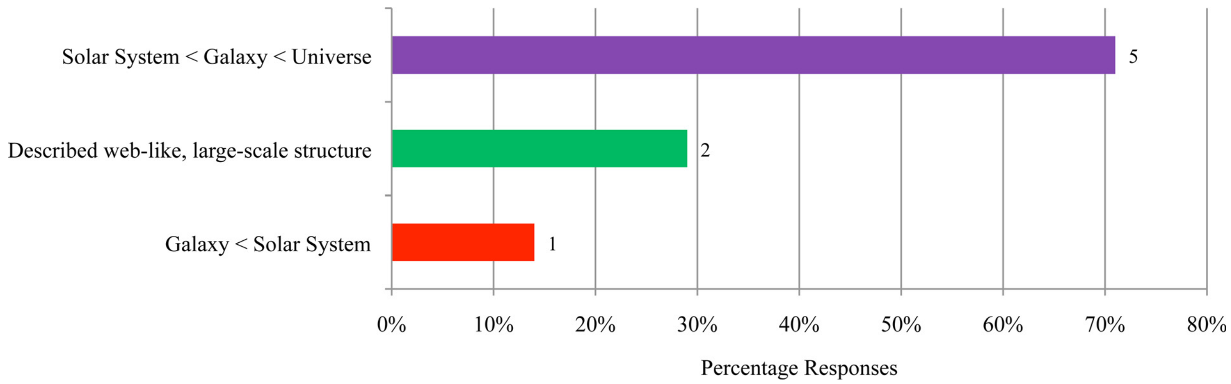


Figure 4.12. Interviews: Relationships. Students were asked to describe the relationships between the terms solar system, galaxy, and universe. $N = 7$. All of these students were interviewed after the introductory material on structure and distance as well as after in-depth instruction about the Solar System. Furthermore, five had received in-depth instruction on the Milky Way Galaxy, four on galaxies, and two on large-scale structure. Additionally, there were four unsolicited correct statements with regard to hierarchy, all made by students after instruction on large-scale structure.

We asked a subset of students in interviews to describe the relationships between the terms solar system, galaxy, and Universe (Figure 4.12, $N = 7$). All of these students were interviewed after the introductory material on structure and distance as well as after in-depth instruction about the Solar System. Furthermore, five had received in-depth instruction on the Milky Way Galaxy, four on galaxies, and two on large-scale structure. Five students (71%) gave correct responses that indicated the hierarchical structure of the Universe, i.e., that a solar system is inside a galaxy, which is contained within the Universe. Two of those students (28%), who were both interviewed following instruction about large-scale structure, also correctly described the web-like large-scale structure of the Universe. Of the remaining two students, one ("C") switched galaxy and solar system and the other ("A") responded, "They are related in the fact that they are all a part of the Universe." It is worth noting that four additional students who were not specifically asked about relationships nonetheless volunteered correct statements about the hierarchical structure during the course of their interviews. All four of these students were interviewed after instruction on large-scale structure.

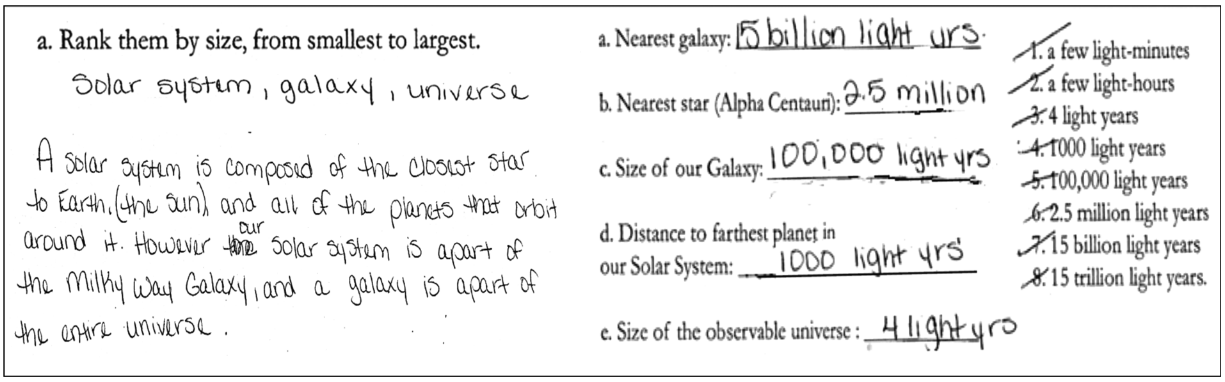
4.5.2 Implications

Data from our precourse homework essays and precourse surveys suggest that students do not have a robust sense of the hierarchical structure of the Universe when entering the classroom. Although many students recognize the importance of orbits, a significant number of students view the Universe as "chaotic, random, or swirly," like a "kaleidoscope."

Structure, particularly the relationships between hierarchical structures, is a theme that appears numerous times throughout the CSU course materials, and students' understanding of the hierarchical relationships deepened as the semester progressed. Although exam questions do not afford a direct comparison, students were able to correctly explain the relationship between the Solar System and the Galaxy more than three-quarters (77%) of the time on the Final Exam. This is higher than the fraction of students who could correctly explain their reasoning on Exam 3 and considerably higher than the number of Correct responses in the reanalyzed CSU subset of the Bailey *et al.* (2012) precourse surveys. In the Bailey *et al.* (2012) precourse surveys conflation of Galaxy and Solar System was the most common error.

In interviews, we saw more nuanced and detailed responses from students who were interviewed later in the course. While this could be due to simply those individual students' level of understanding rather than time, it might also be an indication that additional instruction about the various structural aspects of the Universe and its components reinforces understanding over the course of the semester.

In light of this, instructors should not expect students to fully grasp the hierarchical relationships between solar system, galaxy, and Universe after only an initial introduction early in an ASTRO 101 course (chapter 1 in a typical textbook). However, with further instruction during the course, including more background on these objects and terms, students can come to a deeper understanding of these relationships.


a. Rank them by size, from smallest to largest.

Solar system, galaxy, universe

A solar system is composed of the closest star to Earth. (the sun) and all of the planets that orbit around it. However ~~the~~ our solar system is apart of the Milky Way Galaxy, and a galaxy is apart of the entire universe.

a. Nearest galaxy: 15 billion light urs.

b. Nearest star (Alpha Centauri): 2.5 million

c. Size of our Galaxy: 100,000 light yrs

d. Distance to farthest planet in our Solar System: 1000 light yrs

e. Size of the observable universe: 4 light yrs

1. a few light-minutes
2. a few light-hours
3. 4 light years
4. 1000 light years
5. 100,000 light years
6. 2.5 million light years
7. 15 billion light years
8. 15 trillion light years.

Figure 4.13. This student gives a correct relative ranking and explanation for solar system, galaxy and universe, but does not give absolute scales consistent with the relative rankings and content of the structures.

4.5.3 Implications for students' understanding of the relationship between structure and distance

In addition to looking separately at students' ideas of distance and structure, we also examined themes relevant to both of them together. On exams, we find that students have an easier time with ranking tasks than with tasks involving absolute numbers, both on tasks related to the sizes and distances of individual objects (Subsection 3.2) and tasks related to structure (solar system—galaxy—universe). In interviews, we find that students more frequently cite relative scales than either absolute scales or scale models. Delgado *et al*. (2007) suggest that there is a learning progression, with ranking (or "ordering") being one of the easiest tasks that students pick up and a prerequisite for correctly assigning absolute numbers with units. Students may cue off of a robust relative scale to anchor an absolute scale (Tretter, Jones, and Minogue 2006); however, understanding relative scales does not guarantee the ability to assign correct or consistent absolute scales.

Delgado *et al*. (2007) find consistency between ranking and absolute size tasks about two-thirds (64%) of the time. Figure 4.13 gives an example from our data of a student response to a ranking task and a matching task from Exam 1. This student holds a correct model of relative size and scale with regard to cosmic structures, but the student's choices of absolute scale are inconsistent with the choices of relative size. While we cannot say for certain how this student arrived at their answers, the student is not making robust connections between relative and absolute scales. It is possible that students may just be trying to memorize absolute numbers without reasoning about relationships.

Delgado *et al*. (2007) did not investigate the effect of content knowledge (e.g., the fact that galaxies consist of stars, which themselves form solar systems, etc.). We see that students had an easier time ranking structures (i.e., solar system—galaxy—universe) than lists of individual astronomical objects (e.g., Sun, Jupiter, Andromeda Galaxy), which suggests that content knowledge, or more specifically conceptual understanding of the nature of such objects, is important for their learning.

We suggest that instructors should practice not just ranking, scale model, ratio, and absolute size/distance scale tasks with students, but they should also explicitly practice making connections between these ideas, and tying them to course content.

5. DISCUSSION

In this article, we have described research conducted using multiple data sources over five semesters to catalog students' ideas about astronomical distances and structure throughout the learning process in an active-learning general education astronomy course. In Paper II of this series (Coble *et al*., submitted), we examine student ideas on the composition of the Universe. In Paper III (Trouille *et al*., submitted), we investigate student ideas on the Big Bang Theory, the expansion of the Universe, the age of the Universe, and the timing of cosmological events. Here we have used precourse homework essays, prelab surveys ("pretests"), short and long exam questions, and in-depth interviews to identify student ideas. In a complementary companion study (Bailey *et al*., 2012), we used open-response precourse surveys taken from five institutions (including CSU) over five semesters to identify students' preinstructional ideas in cosmology.

As has been the case in many areas of science, we find that students entering our ASTRO 101 courses bring with them a wide variety of ideas, both aligned with and different from scientific knowledge. Specific to cosmology, we explore a number of common student ideas; our findings include the following:

- Students conflate the term light-year with a measurement of time; this concept appears to be difficult for students to master and retain if not reinforced.
- Relative size and distance scales are easier for students to learn than absolute scales.
- Students tend to initially (incorrectly) think that a star exhibiting a greater parallax shift is farther away.
- When describing objects in the Universe preinstruction, students more often list closer objects, such as Solar System objects, and include more distant objects such as stars and galaxies less frequently.
- Students' concepts of the structure of the Solar System progress only a limited amount from precourse ideas; it is difficult for students to move beyond the correct but basic idea of "planets orbiting the Sun" that they learned in elementary school.
- Students' concepts of galactic structure become more robust throughout the semester; however, the greatest area of difficulty is in visualizing the galactic halo.
- When describing the Universe at the beginning of the course, students most often describe it as dark, with vast, empty spaces, with no edge, and being infinite or big (or both!) in size.
- Students' understanding of the hierarchical nature of structure is weak at the start of a course. Preinstruction, a substantial number of students view the Universe as chaotic, random, or swirly, like a "kaleidoscope" of colors and shapes. However, by the end of the semester, student understanding of the hierarchical nature of structure grows to become fairly robust.

In analyzing our data, we found several themes that crossed over or between our content topics of size and distance and structure:

- First, we see that for concepts that were targeted as themes throughout the semester, like structure, students' understanding continually improved. For concepts that we did not keep coming back to explicitly but with which students likely had some baseline familiarity (or misunderstanding), such as the light-year, students' knowledge increased but later deteriorated.
- Students' understanding of some concepts is impaired by weak math skills, including proportions. These come up in a variety of situations, such as the "half speed of light" problem and the inverse square law, among others. It is encouraging that with continual practice students improve on tasks that use these concepts. For example, understanding proportions from equations is a course theme, and we see that students' ability to answer questions about the inverse square law improves continuously throughout the semester.
- On ranking tasks, students score better on tasks involving hierarchical structures (solar system—galaxy—universe) than on tasks involving individual objects within these structures.

As a result of these findings, we propose the following suggestions for astronomy educators to keep in mind:

- Because students find ranking tasks easier than tasks dealing with absolute scales and might not be consistent in their reasoning between the two, instructors should consider explicitly linking different types of scale activities. This would allow students to practice being internally consistent in their reasoning with size, scale, and distance.
- Student understanding of the structures in the Universe and the hierarchical relationships between them is often conflated preinstruction, but progresses to become fairly robust by the end of a semester. Therefore, instructors should not expect students to have fully formed mental model of structures after an initial unit on size and scope (chapter 1 in a typical ASTRO 101 textbook) but instead should expect this understanding to develop throughout the course.
- Given that undergraduate students often hold tightly to their precollege ideas about the Solar System and that they may enter a course with a myopic view of the Universe as containing little beyond the Solar System, we suggest teaching the hierarchical nature of the Universe, namely, that the Solar System is part of the Galaxy, and that there are many galaxies in the observable Universe, as part of the upper elementary and middle school astronomy curriculum. Given students' difficulties with absolute scales, but greater ease with relative scales, this effort could focus on relative scales, or scale models, which are major themes in elementary and middle school math and science standards.
- Because students have difficulty visualizing the halos of galaxies, we recommend incorporating discussions, activities, and visualizations of halos as part of galactic structure early and often in a course.

By building upon these ideas, we can help students move toward improved understanding of both the processes and outcomes of cosmology, and of science and mathematics more generally. The information in this series of papers is being used as the foundation for creating an innovative cosmology curriculum (Coble *et al*., in preparation) and we would encourage other authors and instructors to use this to inform their curricula as well.

ACKNOWLEDGMENTS

We gratefully thank Virginia Hayes for her help in the data collection and a preliminary analysis of the structure data. Kevin McLin and Anne Metevier provided valuable comments on drafts of this paper. Doug Lombardi gave useful advice on statistics. Mel Sabella assisted with some of the interviews. Of course none of this would be possible without the participation of students in the CSU courses. This work was supported by NASA grant NNX10AC89G and by the Education and Public Outreach program for NASA's Fermi Gamma-ray Space Telescope, the Illinois Space Grant Consortium, NSF CCLI Grant No. #0632563 at Chicago State University.

APPENDIX: DETAILED TABLES FOR SECTIONS 3 AND 4

These tables correspond with the summary tables in the main body of the paper (e.g., Tables 3.1 and A3.1, 3.2 and A3.2, etc.). Appendix tables include the original question, the answer (in bold if short answer), the detailed *N*'s and percentages for different semesters, and the *H*-statistics and *p*-values from KW tests or KS-statistics and *p*-values from KS tests as appropriate.

3. DISTANCE

Table A3.1. Laboratory 1 pretest: Half speed of light.
Question: In the future, when space travel is advanced, you have 3 weeks of vacation time and want to visit the star Sirius, in honor of your favorite Harry Potter character. Sirius is 8.6 light-years away. If spaceships in the future could travel at half the speed of light (much faster than current spaceships), would you be able to make the trip to Sirius and back during your vacation?
Answer: No. If you were traveling at half of the speed of light it would take you 8.6 × 2 = 17.2 years just to get there, so you definitely could not get there and back in 3 weeks.

	Fall 2008 $N = 13$	**Spring 2009 $N = 11$**	**Spring 2010 $N = 19$**	**Fall 2010 $N = 17$**	**Spring 2011 $N = 14$**	**Total $H = 3.16$ $p = 0.53$ $N = 74$**
C	**3 (23%)**	**2 (18%)**	**1 (5%)**	**4 (24%)**	**2 (14%)**	**12 (16%)**
I	2 (15%)	0 (0%)	6 (32%)	4 (24%)	5 (36%)	17 (23%)
P	4 (31%)	3 (27%)	5 (26%)	5 (29%)	1 (7%)	18 (24%)
W	3 (23%)	6 (55%)	5 (26%)	4 (24%)	3 (21%)	21 (28%)
T	1 (8%)	0 (0%)	2 (11%)	0 (0%)	2 (14%)	5 (7%)
NS	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	0 (0%)	0 (0%)	0 (0%)	0 (0%)	1 (7%)	1 (1%)

Table A3.2. Laboratory 1 pretest: Speed and scale.

Question: What are some other astronomical objects you would be able to visit during your vacation? (Again, in a future spaceship that could travel at half the speed of light.)

Answer: Most objects in our Solar System are within a few light-minutes or light-hours away, so you could get to any of those during your futuristic vacation.

	Fall 2008 $N = 13$	**Spring 2009 $N = 11$**	**Spring 2010 $N = 19$**	**Fall 2010 $N = 17$**	**Spring 2011 $N = 14$**	**Total $H = 7.25$ $p = 0.12$ $N = 74$**
a. Content						
C	**5 (38%)**	**9 (82%)**	**8 (42%)**	**8 (47%)**	**4 (29%)**	**34 (46%)**
I	0 (0%)	0 (0%)	0 (0%)	0 (0%)	2 (14%)	2 (3%)
P	4 (31%)	2 (18%)	7 (37%)	4 (24%)	3 (21%)	20 (27%)
W	3 (23%)	0 (0%)	3 (16%)	4 (24%)	1 (7%)	10 (14%)
T	1 (8%)	0 (0%)	1 (5%)	0 (0%)	0 (0%)	2 (3%)
NS	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	0 (0%)	0 (0%)	0 (0%)	1 (6%)	4 (28%)	5 (7%)
b. Reasoning						
C	**1 (8%)**	**1 (9%)**	**0 (0%)**	**1 (6%)**	**1 (7%)**	**4 (5%)**
I	0 (0%)	1 (9%)	4 (21%)	3 (18%)	2 (14%)	10 (14%)
P	1 (8%)	0 (0%)	2 (11%)	3 (18%)	0 (0%)	5 (7%)
W	0 (0%)	1 (9%)	1 (5%)	0 (0%)	0 (0%)	2 (3%)
T	0 (0%)	2 (18%)	0 (0%)	0 (0%)	1 (7%)	3 (4%)
NS	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	11 (85%)	6 (55%)	12 (63%)	10 (59%)	10 (71%)	49 (67%)

Table A3.3. Exam 1: Half speed of light (FIB).

Question:

v1: The star Vega is 25 light-years away. If you were in a spaceship that could travel at half the speed of light, the amount of time it would take you reach Vega is **50 years**. (Be specific, use a number.)

v2: The star Sirius is 9 light-years away. If you were in a spaceship traveling at half of the speed of light, the amount of time it would take you reach Sirius is **18 years**. (Be specific, use a number.)

v3: The Whirlpool galaxy is about 30 million light-years away. If you were in a spaceship that could travel at half of the speed of light, the amount of time it would take you reach the Whirlpool galaxy is **60 million years**. (Be specific, use a number.)

	Fall 2008 (v1) $N = 12$	**Spring 2009 (v2) $N = 9$**	**Spring 2010 (v1) $N = 17$**	**Fall 2010 (v3) $N = 14$**	**Spring 2011 (v3) $N = 13$**	**Total $H = 2.24$ $p = 0.69$ $N = 65$**
C	**5 (42%)**	**3 (33%)**	**6 (35%)**	**7 (50%)**	**5 (38%)**	**26 (40%)**
I	0 (0%)	0 (0%)	1 (6%)	0 (0%)	2 (15%)	3 (5%)
P	0 (0%)	0 (0%)	2 (12%)	0 (0%)	2 (15%)	4 (6%)
W	7 (58%)	4 (44%)	8 (47%)	6 (43%)	3 (23%)	28 (43%)
T	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NS	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	0 (0%)	2 (22%)	0 (0%)	1 (7%)	1 (8%)	4 (6%)

Table A3.4. Exam 1: Speed of light (FIB).

a. Travel time

Question: Our Galaxy is about 100,000 light-years across. If you could travel at the speed of light, how long would it take you to go across our Galaxy? **100 000 years**.

b. Lookback time

Question: The Whirlpool galaxy is about 30 million light-years away. If you went outside and looked at the Whirlpool, how long ago did the light that just arrived at your eye leave its home? **30 million years**.

a. Travel time

	Fall 2008 $N = 12$	**Spring 2009 $N = 9$**	**Spring 2010 $N = 17$**	**Fall 2010**	**Spring 2011**	**Total $H = 1.09$ $p = 0.58$ $N = 38$**
C	**8 (67%)**	**5 (56%)**	**12 (71%)**	—	—	**25 (66%)**
I	1 (8%)	0 (0%)	1 (6%)	—	—	2 (5%)
P	0 (0%)	2 (22%)	0 (0%)	—	—	2 (5%)
W	3 (25%)	1 (11%)	4 (24%)	—	—	8 (21%)
T	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)
NS	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)
NR	0 (0%)	1 (11%)	0 (0%)	—	—	1 (3%)

b. Lookback time

	Fall 2008	**Spring 2009**	**Spring 2010**	**Fall 2010 $N = 14$**	**Spring 2011 $N = 13$**	**Total *KS-stat* = 0.15 $p = 0.98$ $N = 27$**
C	—	—	—	**10 (71%)**	**7 (54%)**	**17 (63%)**
I	—	—	—	0 (0%)	2 (15%)	2 (7%)
P	—	—	—	0 (0%)	2 (15%)	2 (7%)
W	—	—	—	3 (21%)	2 (15%)	5 (19%)
T	—	—	—	0 (0%)	0 (0%)	0 (0%)
NS	—	—	—	0 (0%)	0 (0%)	0 (0%)
NR	—	—	—	1 (7%)	0 (0%)	1 (4%)

Table A3.5. Exam 1: Definition of light-year (MC).

Question

v1: A television advertisement claiming that a product is light-years ahead of its time does not make sense because

- a. light-years can only be used to talk about light
- b. it doesn't specify the number of light-years
- **c. it uses "light-years" to talk about time, but a light-year is a unit of distance**
- d. a light-year is an astronomically large unit, so a product could not possibly be so advanced

v2: When we look at an object that is 1000 light-years away we see it

- a. as it is right now, but it appears 1000 times dimmer
- b. looking just the same as our ancestors would have seen it 1000 years ago
- c. as it was 1000 light-years ago
- **d. as it was 1000 years ago**

a. Question v1

	Fall 2008	**Spring 2009 $N = 9$**	**Spring 2010**	**Fall 2010 $N = 14$**	**Spring 2011 $N = 13$**	**Total $H = 0.14$ $p = 0.93$ $N = 36$**
A	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)
B	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)
C	—	**9 (100%)**	—	**14 (100%)**	**12 (92%)**	**35 (97%)**
D	—	0 (0%)	—	0 (0%)	1 (8%)	1 (3%)
NR	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)

b. Question v2

	Fall 2008 $N = 12$	**Spring 2009**	**Spring 2010 $N = 17$**	**Fall 2010**	**Spring 2011**	**Total *KS-stat* = 0.35 $p = 0.25$ $N = 29$**
A	1 (8%)	—	1 (6%)	—	—	2 (7%)
B	1 (8%)	—	1 (6%)	—	—	2 (7%)
C	2 (17%)	—	1 (6%)	—	—	3 (10%)
D	**8 (67%)**	—	**13 (76%)**	—	—	**21 (72%)**
NR	0 (0%)	—	1 (6%)	—	—	1 (3%)

Table A3.6. Final exam: Half speed of light (FIB).

Question:

v1: The star Procyon is about 11 light-years away. If you were in a spaceship that could travel at half of the speed of light, how long would it take you to get to Procyon? **22 years**.

v2: The star Sirius is 9 light-years away. If you were in a spaceship that could travel at half of the speed of light, the amount of time it would take you reach Sirius is **18 years**. (Be specific, use a number.)

	Fall 2008	**Spring 2009 (v1) $N = 9$**	**Spring 2010 (v2) $N = 13$**	**Fall 2010 (v1) $N = 13$**	**Spring 2011 (v1) $N = 13$**	**Total $H = 6.06$ $p = 0.11$ $N = 48$**
C	—	**5 (56%)**	**3 (23%)**	**5 (38%)**	**3 (23%)**	**16 (33%)**
I	—	0 (0%)	0 (0%)	1 (8%)	1 (8%)	2 (4%)
P	—	2 (22%)	5 (38%)	0 (0%)	7 (54%)	14 (29%)
W	—	2 (22%)	4 (31%)	6 (46%)	1 (8%)	13 (27%)
T	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NS	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	—	0 (0%)	1 (8%)	1 (8%)	1 (8%)	3 (6%)

Table A3.7. Final Exam: Speed of light lookback time (FIB).

Question: The Andromeda galaxy is about 2.5 million light-years away. If you went outside and looked at Andromeda, how long ago did the light that just arrived at your eye leave its home? **2.5 million years**.

	Fall 2008	**Spring 2009**	**Spring 2010 $N = 13$**	**Fall 2010 $N = 13$**	**Spring 2011 $N = 13$**	**TOTAL $H = 0.57$ $p = 0.75$ $N = 39$**
C	—	—	**8 (62%)**	**10 (77%)**	**8 (62%)**	**26 (67%)**
I	—	—	1 (8%)	0 (0%)	0 (0%)	1 (3%)
P	—	—	3 (23%)	1 (8%)	5 (38%)	9 (23%)
W	—	—	1 (8%)	1 (8%)	0 (0%)	2 (5%)
T	—	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NS	—	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	—	—	0 (0%)	1 (8%)	0 (0%)	1 (3%)

Table A3.8. Final Exam: Definition of light-year (MC, v2).

Question: When we look at an object that is 1000 light-years away we see it

a. as it was 1000 light-years ago
b. as it was 1000 years ago
c. looking just the same as our ancestors would have seen it 1000 years ago
d. as it is right now, but it appears 1000 times dimmer

	Fall 2008 $N = 10$	**Spring 2009 $N = 9$**	**Spring 2010 $N = 13$**	**Fall 2010 $N = 13$**	**Spring 2011 $N = 13$**	**Total $H = 0.57$ $p = 0.97$ $N = 58$**
A	0 (0%)	1 (11%)	0 (0%)	1 (8%)	0 (0%)	2 (3%)
B	**10 (100%)**	**7 (78%)**	**11 (85%)**	**11 (85%)**	**12 (92%)**	**51 (88%)**
C	0 (0%)	0 (0%)	1 (8%)	0 (0%)	0 (0%)	1 (2%)
D	0 (0%)	1 (11%)	1 (8%)	1 (8%)	1 (8%)	4 (7%)
NR	0 (0%)	0 (0%)	1 (6%)	0 (0%)	0 (0%)	0 (0%)

Table A3.9. Exam 1: Absolute scales (Matching).

Question: Match the distance to each of the following to the closest answer.

a. Nearest galaxy: **(6) 2.5 million light-years**
b. Nearest star (Alpha Centauri): **(3) 4 light-years**
c. Size of our Galaxy: **(5) 100 000 light-years**
d. Distance to farthest planet in our Solar System: **(2) few light-hours**
e. Size of the observable universe: **(7) 15 billion light-years**

1. a few light-minutes
2. a few light-hours
3. 4 light-years
4. 1000 light-years
5. 100 000 light-years
6. 2.5 million light-years
7. 15 billion light-years
8. 15 trillion light-years

	Fall 2008 $N = 12$	**Spring 2009 $N = 9$**	**Spring 2010 $N = 17$**	**Fall 2010 $N = 14$**	**Spring 2011 $N = 13$**	**Total $H = 1.85$ $p = 0.76$ $N = 65$**
5	**1 (8%)**	**3 (33%)**	**4 (24%)**	**4 (29%)**	**2 (15%)**	**14 (22%)**
4	2 (17%)	1 (11%)	4 (24%)	1 (7%)	0 (0%)	8 (12%)
3	3 (25%)	1 (11%)	3 (18%)	1 (7%)	2 (15%)	10 (15%)
2	4 (33%)	2 (22%)	1 (6%)	3 (21%)	4 (31%)	14 (22%)
1	1 (8%)	2 (22%)	2 (12%)	3 (21%)	5 (38%)	13 (20%)
0	1 (8%)	0 (0%)	3 (18%)	1 (7%)	0 (0%)	5 (8%)
NR	0 (0%)	0 (0%)	0 (0%)	1 (7%)	0 (0%)	1 (2%)

Table A3.10. Exam 1: Object distance ranking.

Question: Consider these five objects:

a. The Andromeda Galaxy
b. The Pleiades Star Cluster
c. The Hubble Space Telescope
d. Jupiter
e. The Sun

Rank them by: distance away from Earth (from closest to farthest): **C E D B A**

	Fall 2008	**Spring 2009 $N = 9$**	**Spring 2010**	**Fall 2010 $N = 14$**	**Spring 2011 $N = 13$**	**Total $H = 2.10$ $p = 0.35$ $N = 36$**
5	—	**7 (78%)**	—	**7 (50%)**	**6 (46%)**	**20 (56%)**
4	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)
3	—	2 (22%)	—	4 (29%)	7 (54%)	13 (36%)
2	—	0 (0%)	—	1 (7%)	0 (0%)	1 (3%)
1	—	0 (0%)	—	2 (14%)	0 (0%)	2 (6%)
0	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)
NR	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)

Table A3.11. Exam 1: Object size ranking.

Question: Consider these seven objects: (1) The Moon, (2) The Andromeda Galaxy, (3) The Coma Cluster of Galaxies, (4) The Hubble Space Telescope, (5) The Pleiades Star Cluster, (6) Saturn, and (7) The Sun. Rank them by size.

Answer: 4 1 6 7 5 2 3

(4) The Hubble Space Telescope: man-made, about the size of a bus
(1) The Moon: about 1/4 the size of the Earth
(6) Saturn: about 10 times bigger across than Earth
(7) The Sun: about 100 times bigger across than Earth, (also it's a star, so it's bigger than planets)
(5) Pleiades Star Cluster: contains about 1000 stars-**>** bigger than the Sun but smaller than a galaxy
(2) Andromeda Galaxy: a galaxy contains billions of stars
(3) Coma Cluster of Galaxies: contains thousands of galaxies

	Fall 2008 $N = 12$	**Spring 2009**	**Spring 2010 $N = 17$**	**Fall 2010**	**Spring 2011**	**Total $H = 0.21$ $p = 0.78$ $N = 29$**
a. Ranking						
7	**7 (58%)**	—	**9 (53%)**	—	—	**16 (55%)**
6	0 (0%)	—	1 (6%)	—	—	1 (3%)
5	2 (17%)	—	3 (18%)	—	—	5 (17%)
4	2 (17%)	—	2 (12%)	—	—	4 (14%)
3	1 (8%)	—	0 (0%)	—	—	1 (3%)
2	0 (0%)	—	1 (6%)	—	—	1 (3%)
1	0 (0%)	—	1 (6%)	—	—	1 (3%)
0	0 (0%)	—	0 (0%)	—	—	0 (0%)
NR	0 (0%)	—	0 (0%)	—	—	0 (0%)
b. Reasoning						
C	**4 (33%)**	—	**3 (18%)**	—	—	**7 (24%)**
I	3 (25%)	—	6 (35%)	—	—	9 (31%)
P	3 (25%)	—	3 (18%)	—	—	6 (21%)
W	2 (17%)	—	2 (12%)	—	—	4 (14%)
T	0 (0%)	—	3 (18%)	—	—	3 (10%)
NS	0 (0%)	—	0 (0%)	—	—	0 (0%)
NR	0 (0%)	—	0 (0%)	—	—	0 (0%)

Table A3.12. Final exam: Object distance ranking.

Question: Consider these five objects:

a. The Andromeda Galaxy
b. The Pleiades Star Cluster
c. The Hubble Space Telescope
d. Jupiter
e. The Sun

Rank them by: distance away from Earth (from closest to farthest): **C E D B A**

	Fall 2008 $N = 10$	**Spring 2009 $N = 9$**	**Spring 2010 $N = 13$**	**Fall 2010**	**Spring 2011**	**Total $H = 1.69$ $p = 0.43$ $N = 32$**
5	**4 (40%)**	**7 (78%)**	**8 (62%)**	—	—	**19 (59%)**
4	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)
3	4 (40%)	1 (11%)	2 (15%)	—	—	7 (22%)
2	0 (0%)	0 (0%)	2 (15%)	—	—	2 (6%)
1	2 (20%)	1 (11%)	1 (8%)	—	—	4 (13%)
0	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)
NR	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)

Table A3.13. Final exam: Object size ranking task.

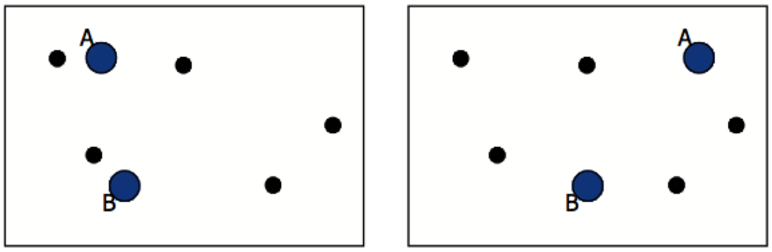
Question: Consider these five objects:

a. The Andromeda Galaxy
b. The Hubble Space Telescope
c. Jupiter
d. The Pleiades Star Cluster
e. The Sun

Rank them by size (from smallest to largest): **B C E D A**

	Fall 2008	**Spring 2009**	**Spring 2010**	**Fall 2010 $N = 13$**	**Spring 2011 $N = 13$**	**Total $KS\text{-}stat = 0.13$ $p = 0.99$ $N = 26$**
5	—	—	—	**8 (62%)**	**9 (69%)**	**17 (65%)**
4	—	—	—	0 (0%)	0 (0%)	0 (0%)
3	—	—	—	2 (15%)	3 (23%)	5 (19%)
2	—	—	—	1 (8%)	1 (8%)	2 (8%)
1	—	—	—	2 (15%)	0 (0%)	2 (8%)
0	—	—	—	0 (0%)	0 (0%)	0 (0%)
NR	—	—	—	0 (0%)	0 (0%)	0 (0%)

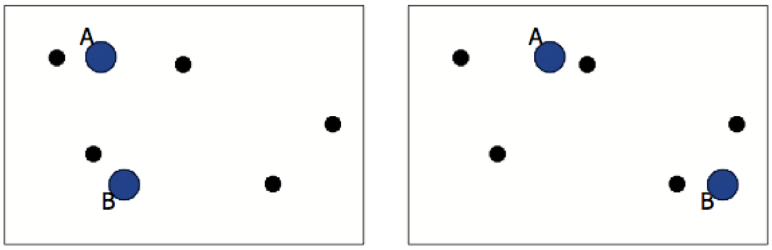
Table A3.14. Laboratory 9 pretest: Parallax measurements.
Question: Parallax. The following two pictures were taken six months apart. Which star is farther away, A or B?[a]
Answer: B is farther away



	Fall 2008 $N = 9$	**Spring 2009 $N = 8$**	**Spring 2010 $N = 8$**	**Fall 2010 $N = 11$**	**Spring 2011**	**Total $H = 4.44$ $p = 0.22$ $N = 36$**
A	4 (44%)	5 (63%)	6 (75%)	4 (36%)	—	19 (53%)
B	**4 (44%)**	**2 (25%)**	**2 (25%)**	**2 (18%)**	—	**10 (28%)**
SAME	1 (11%)	0 (0%)	0 (0%)	5 (45%)	—	6 (17%)
NR	0 (0%)	1 (13%)	0 (0%)	0 (0%)	—	1 (3%)

[a] Student responses to a second question, “How do you know? Are given in Figure 3.8.

Table A3.15a. Exam 3: Parallax measurements (FIB).
Question:
v1: Parallax. The following two pictures were taken six months apart. Star **A** is farther away.
v2: Parallax. The following two pictures were taken six months apart. Star **B** is closer.



	Fall 2008 (v1) $N = 9$	**Spring 2009 (v1) $N = 9$**	**Spring 2010 (v2) $N = 12$**	**Fall 2010 (v2) $N = 13$**	**Spring 2011 (v2) $N = 13$**	**Total $H = 1.67$ $p = 0.80$ $N = 56$**
C	**6 (67%)**	**6 (67%)**	**6 (50%)**	**6 (46%)**	**5 (38%)**	**29 (51%)**
W	3 (33%)	2 (22%)	6 (50%)	6 (46%)	8 (62%)	25 (45%)
NR	0 (0%)	1 (11%)	0 (0%)	1 (8%)	0 (0%)	2 (4%)

Table A3.15b. Exam 3: Parallax, numerical reasoning (MC, v1).
Question:
v1: Star X is known to be 10 parsecs away from us and star Y is 50 parsecs away. Which star has the greater parallax shift (angle)?

- **a. Star X**
- b. Star Y
- c. There is insufficient information to determine this.
- d. Neither—their parallax angles are the same.

v2: You observe two stars over the course of a year (or more) and find that both stars have measurable parallax angles. Star X has a parallax angle of 1 arc sec. Star Y has a parallax angle of 2 arc sec. Which star is closer?

- **a. Star X**
- b. Star Y
- c. There is insufficient information to determine this.

	Fall 2008 (v1) *N* = 9	**Spring 2009**	**Spring 2010**	**Fall 2010 (v2) *N* = 13**	**Spring 2011 (v2) *N* = 13**	**Total *H* = 0.65 *p* = 0.72 *N* = 35**
A	**5 (56%)**	—	—	**7 (54%)**	**5 (38%)**	**17 (49%)**
B	3 (33%)	—	—	3 (23%)	5 (38%)	11 (31%)
C	1 (11%)	—	—	3 (23%)	3 (23%)	7 (20%)
D	0 (0%)	—	—	—	—	0 (0%)
NR	0 (0%)	—	—	0 (0%)	0 (0%)	0 (0%)

Table A3.15c. Exam 3: Parallax, numerical reasoning (MC, v2).
Question: Which of the following stars is closest to us?

- a. Procyon (parallax angle = 0.29 arc sec)
- b. Ross 780 (parallax angle = 0.21 arc sec)
- c. Regulus (parallax angle = 0.04 arc sec)
- **d. Sirius (parallax angle = 0.38 arc sec)**

	Fall 2008	**Spring 2009 *N* = 9**	**Spring 2010 *N* = 12**	**Fall 2010**	**Spring 2011**	**Total *H* = 0.01 *p* = 0.92 *N* = 21**
A	—	0 (0%)	0 (0%)	—	—	0 (0%)
B	—	0 (0%)	0 (0%)	—	—	0 (0%)
C	—	4 (44%)	5 (42%)	—	—	9 (43%)
D	—	**5 (56%)**	**7 (58%)**	—	—	**12 (57%)**
NR	—	0 (0%)	0 (0%)	—	—	0 (0%)

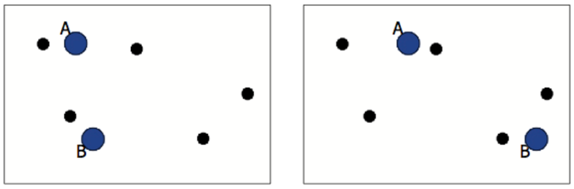
Table A3.16. Final exam: Parallax measurements (FIB).

Question:

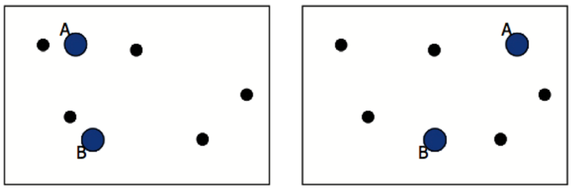
v1: Parallax. The following two pictures were taken six months apart. Star **B** is closer.

v2: Parallax. The following two pictures were taken six months apart. Star **A** is farther away.

v3: Parallax. The following two pictures were taken six months apart. Star **B** is farther away.



(Figure for v1, v2)



(Figure for v3)

	Fall 2008 (v1) $N = 10$	**Spring 2009 (v1)** $N = 9$	**Spring 2010 (v2)** $N = 13$	**Fall 2010 (v3)** $N = 13$	**Spring 2011**	**Total** $H = 4.23$ $p = 0.24$ $N = 45$
C	**8 (80%)**	**6 (67%)**	**8 (62%)**	**4 (31%)**	—	**26 (58%)**
W	2 (20%)	3 (33%)	4 (31%)	9 (69%)	—	18 (40%)
NR	0 (0%)	0 (0%)	1 (8%)	0 (0%)	—	1 (2%)

Table A3.17a. Laboratory 9 pretest: Inverse square law.

Question: Star C and Star D have the same luminosity (inherent brightness), but Star C is 5 times farther away than Star D. If the intensity of light is inversely proportional to distance squared, how does the apparent brightness of star D compare to that of star C. Be specific (give a number) and explain your reasoning.

Answer: If star C is 5 times farther away, it will be $5^2 = 25$ times dimmer than star D. (Also acceptable: star D will be 25 times brighter than star C.)

	Fall 2008 $N = 9$	**Spring 2009** $N = 8$	**Spring 2010** $N = 8$	**Fall 2010** $N = 11$	**Spring 2011**	**Total** $H = 0.79$ $p = 0.37$ $N = 36$
C	**2 (22%)**	**2 (25%)**	**2 (25%)**	**3 (33%)**	—	**10 (28%)**
I	1 (11%)	2 (25%)	1 (13%)	1 (11%)	—	5 (14%)
P	4 (44%)	2 (25%)	3 (38%)	1 (11%)	—	11 (31%)
W	2 (22%)	2 (25%)	1 (13%)	2 (22%)	—	8 (22%)
T	0 (0%)	0 (0%)	1 (13%)	0 (0%)	—	1 (3%)
NS	0 (0%)	0 (0%)	0 (0%)	0 (0%)	—	0 (0%)
NR	0 (0%)	0 (0%)	0 (0%)	1 (11%)	—	1 (3%)

Table A3.17b. Exam 3: Inverse square law (FIB).
Question:

v1: Stars A and B are both Cepheid variable stars. Star A is 6 times farther away than star B. The brightness of star A will be **$6^2 = 36$ times dimmer** compared to B. (Use a number.)

v2: Stars A and B both have the same spectral class. Star A is 7 times farther away than star B. The brightness of star A will be **$7^2 = 49$ times dimmer** compared to B. (Give a number.)

v3: Supernova A is 10 times farther away than supernova B. The apparent brightness of supernova A will be **$10^2 = 100$ times dimmer** compared to B. (Give a number.)

v4: Supernova A is 10 times farther away than supernova B. The apparent brightness of supernova A will be **$10^2 = 100$** times dimmer than B. (Give a number.)

	Fall 2008 (v1) $N = 9$	**Spring 2009 (v4) $N = 9$**	**Spring 2010 (v2) $N = 12$**	**Fall 2010 (v3) $N = 13$**	**Spring 2011 (v3) $N = 13$**	**Total $H = 3.82$ $p = 0.43$ $N = 56$**
C	**2 (22%)**	**6 (67%)**	**4 (33%)**	**5 (38%)**	**6 (46%)**	**23 (41%)**
I	2 (22%)	0 (0%)	0 (0%)	4 (31%)	3 (23%)	9 (16%)
P	3 (33%)	0 (0%)	3 (25%)	3 (23%)	3 (23%)	12 (21%)
W	2 (22%)	3 (33%)	5 (42%)	1 (8%)	1 (8%)	12 (21%)
T	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NS	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Table A3.17c. Final exam: Inverse square law (FIB).
Question:

v1: Stars A and B are both G-type stars. Star A is 3 times closer than star B. The brightness of star A will be **$3^2 = 9$ times brighter** compared to B. (Use a number.)

v2: Stars A and B are both Cepheid variable stars. Star A is 5 times closer than star B. The brightness of star A will be **$5^2 = 25$ times brighter** compared to B. (Use a number.)

v3: Stars A and B are both Cepheid variable stars. Star A is 6 times closer than star B. The brightness of star A will be **$6^2 = 36$ times brighter** compared to B. (Use a number.)

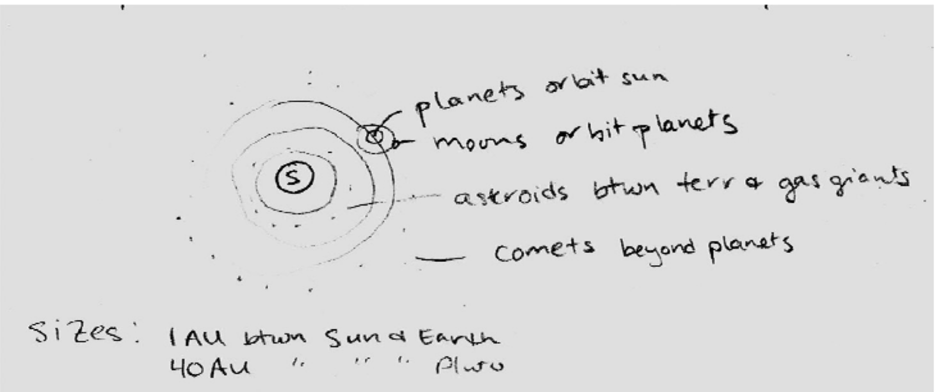
	Fall 2008 (v1) $N = 10$	**Spring 2009 (v2) $N = 9$**	**Spring 2010 (v3) $N = 13$**	**Fall 2010**	**Spring 2011**	**Total $H = 2.86$ $p = 0.24$ $N = 32$**
C	**4 (40%)**	**7 (78%)**	**8 (62%)**	—	—	**19 (59%)**
I	2 (20%)	2(22%)	1 (8%)	—	—	5 (16%)
P	3 (30%)	0 (0%)	1 (8%)	—	—	4 (13%)
W	1 (10%)	0 (0%)	3 (23%)	—	—	4 (13%)
T	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)
NS	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)
NR	0 (0%)	0 (0%)	0 (0%)	—	—	0 (0%)

4. STRUCTURE

Table A4.1. Precourse survey reanalysis: Solar System.
$N = 17$ written responses were collected and analyzed.

C	I	P	W	T	NS	NR
0 (0%)	12 (71%)	2 (12%)	2 (12%)	0 (0%)	0 (0%)	1 (6%)

Table A4.2. Final exam: Solar System structure (essay).
Question: Draw a diagram of our solar system. Label 3 major types of components of the solar system and give the approximate size of at least one of those components.
Answer:



	Fall 2008	Spring 2009 $N = 9$	Spring 2010 $N = 13$	Fall 2010 $N = 13$	Spring 2011 $N = 13$	Total $H = 1.11$ $p = 0.77$ $N = 48$
C	—	**1 (11%)**	**1 (8%)**	**4 (31%)**	**2 (15%)**	**8 (17%)**
I	—	3 (33%)	5 (39%)	5 (39%)	3 (23%)	18 (38%)
P	—	5 (56%)	5 (39%)	4 (31%)	6 (46%)	18 (38%)
W	—	0 (0%)	1 (8%)	0 (%)	2 (15%)	3 (6%)
T	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NS	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	—	0 (0%)	1 (8%)	0 (0%)	0 (0%)	1 (2%)

Table A4.3. Precourse survey reanalysis: galaxy.
$N = 17$ written responses were collected and analyzed.

C	I	P	W	T	NS	NR
1 (6%)	8 (47%)	3 (18%)	1 (6%)	0 (0%)	0 (0%)	4 (24%)

Table A4.4. Exam 3: Number of Stars in Galaxy (MC).

Question: The number of stars in the Milky Way is approximately?

a. a few hundred million
b. a few hundred thousand
c. a few hundred billion
d. a few hundred

	Fall 2008	**Spring 2009**	**Spring 2010**	**Fall 2010 $N = 13$**	**Spring 2011 $N = 13$**	**Total *KS-stat* = 0.07 $p = 0.99$ $N = 26$**
A	—	—	—	1 (8%)	2 (15%)	3 (12%)
B	—	—	—	1 (8%)	1 (8%)	2 (8%)
C	—	—	—	**11 (85%)**	**10 (77%)**	**21 (80%)**
D	—	—	—	0 (0%)	0 (0%)	0 (0%)
NR	—	—	—	0 (0%)	0 (0%)	0 (0%)

Table A4.5. Exam 3: Galaxy types (MC).

Question: Which of the following is NOT one of the three major categories of galaxies?

a. Globular Galaxies
b. Spiral Galaxies
c. Elliptical Galaxies
d. Irregular Galaxies

	Fall 2008 $N = 9$	**Spring 2009**	**Spring 2010 $N = 12$**	**Fall 2010 $N = 13$**	**Spring 2011 $N = 13$**	**Total $H = 0.37$ $p = 0.95$ $N = 47$**
A	**9 (100%)**	—	**11 (92%)**	**12 (92%)**	**11 (85%)**	**43 (91%)**
B	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)
C	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)
D	0 (0%)	—	1 (8%)	1 (8%)	2 (15%)	4 (9%)
NR	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)

Table A4.6. Exam 3: Galaxy types, Spiral Galaxies (FIB).
a. Spiral Galaxies:
Question: Use the three images of galaxies shown below (A, B, and C) to answer the following question.
v1: In which of the galaxies shown would you expect to see many bright blue stars? *Choose all that apply*.
v2: In which of the galaxies shown would you expect to see regions of abundant gas and dust? *Choose all that apply*.
v3: In which of the galaxies shown would you expect to see mostly blue stars? *Choose all that apply*.
v4: In which of the galaxies shown below would you expect to see young, hot, blue stars? *Choose all that apply*.
Answer: A, C.
b. Elliptical Galaxies:
Question: Use the three images of galaxies shown below (A, B, and C) to answer the following question.
In which of the galaxies would you expect to see mostly red stars? *Choose all that apply*.
Answer: B.


A


B


C

a. Spirals

	Fall 2008 (v1) $N = 9$	Spring 2009 (v2) $N = 9$	Spring 2010	Fall 2010 (v3) $N = 13$	Spring 2011 (v4) $N = 13$	Total $H = 0.78$ $p = 0.85$ $N = 44$
C	**4 (44%)**	**6 (67%)**	—	**7 (54%)**	**4 (31%)**	**21 (48%)**
I	4 (44%)	1 (11%)	—	2 (15%)	8 (62%)	15 (34%)
P	0 (0%)	2 (22%)	—	3 (23%)	0 (0%)	5 (11%)
W	1 (11%)	0 (%)	—	0 (0%)	1 (8%)	2 (5%)
T	0 (0%)	0 (0%)	—	1 (8%)	0 (0%)	1 (2%)

b. Ellipticals

	Fall 2008	Spring 2009	Spring 2010 $N = 12$	Fall 2010	Spring 2011	Total $N = 12$
C	—	—	**4 (33%)**	—	—	**4 (33%)**
P	—	—	3 (25%)	—	—	3 (25%)
W	—	—	5 (42%)	—	—	5 (42%)
NR	—	—	0 (0%)	—	—	0 (0%)

Table A4.7. Exam 3: Galaxy structure.

Question:

v1: On the following artist's conception of our Galaxy, label the bulge, disk, and halo, and draw a scale bar for size.

v2: Draw a diagram of the Milky Way, both edge on and face on. Be sure to label 3 primary components of the galaxy and the approximate size of at least one of those components.

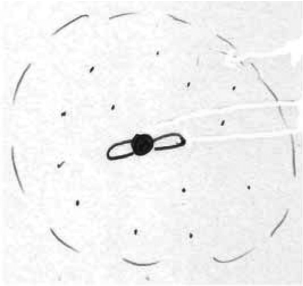
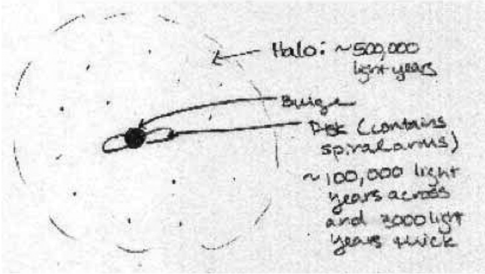


Figure for v1



Answer

	Fall 2008	Spring 2009	Spring 2010	Fall 2010 *N* = 13	Spring 2011 *N* = 13	Total *KS-stat* = 0.16 *p* = 0.98 *N* = 26
a. Label (v1, FIB)						
C	—	—	—	**2 (15%)**	**4 (31%)**	**6 (23%)**
I	—	—	—	6 (46%)	3 (23%)	9 (35%)
P	—	—	—	5 (39%)	5 (39%)	10 (39%)
W	—	—	—	0 (0%)	1 (8%)	1 (4%)
T	—	—	—	0 (0%)	0 (0%)	0 (0%)
NS	—	—	—	0 (0%)	0 (0%)	0 (0%)
NR	—	—	—	0 (0%)	0 (0%)	0 (0%)

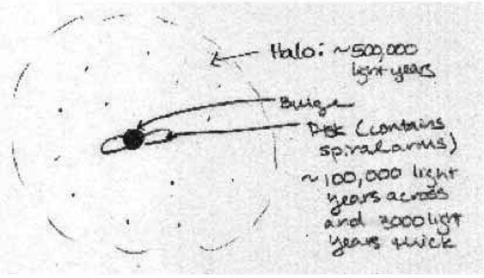
b. Draw and label (v2, Essay)

	Fall 2008 *N* = 9	Spring 2009	Spring 2010 *N* = 12	Fall 2010	Spring 2011	Total *KS-stat* = 0.21 *p* = 0.86 *N* = 21
C	**0 (0%)**	—	**2 (17%)**	—	—	**2 (10%)**
I	1 (11%)	—	2 (17%)	—	—	3 (14%)
P	8 (88%)	—	4 (33%)	—	—	12 (57%)
W	0 (0%)	—	0 (0%)	—	—	0 (0%)
T	0 (0%)	—	2 (17%)	—	—	2 (10%)
NS	0 (0%)	—	0 (0%)	—	—	0 (0%)
NR	0 (0%)	—	2 (17%)	—	—	2 (10%)

Table A4.8. Final exam: Galaxy structure, draw and label (essay).

Question: Draw a diagram of the Milky Way Galaxy. Label 3 major types of components of the Galaxy and give the approximate size of at least one of those components.

Answer:



	Fall 2008	Spring 2009 $N = 9$	Spring 2010 $N = 13$	Fall 2010 $N = 13$	Spring 2011 $N = 13$	Total $H = 0.74$ $p = 0.86$ $N = 48$
C	—	**2 (22%)**	**4 (31%)**	**5 (39%)**	**3 (23%)**	**14 (29%)**
I	—	2 (22%)	4 (31%)	2 (15%)	4 (31%)	12 (25%)
P	—	5 (56%)	3 (23%)	6 (46%)	4 (31%)	18 (38%)
W	—	0 (0%)	0 (0%)	0 (0%)	2 (15%)	2 (4%)
T	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NS	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	—	0 (0%)	2 (15%)	0 (0%)	0 (0%)	2 (4%)

Table A4.9. Final exam: Definition of galaxy (MC).

Question: A typical galaxy is a

a. nearby object orbiting a planet
b. large, glowing ball of gas powered by nuclear energy
c. relatively small, icy object orbiting a star
d. collection of a few hundred million to a trillion or more stars, bound together by gravity
e. system consisting of one or a few stars orbited by planets, moons, and smaller objects

	Fall 2008 $N = 10$	Spring 2009 $N = 9$	Spring 2010 $N = 13$	Fall 2010 $N = 13$	Spring 2011 $N = 13$	Total $H = 1.25$ $p = 0.87$ $N = 58$
A	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
B	0 (0%)	0 (0%)	1 (8%)	0 (0%)	0 (0%)	1 (2%)
C	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
D	**9 (90%)**	**7 (78%)**	**7 (54%)**	**12 (92%)**	**10 (77%)**	**45 (78%)**
E	1 (10%)	2 (22%)	5 (38%)	0 (0%)	3 (23%)	11 (19%)
NR	0 (0%)	0 (0%)	0 (0%)	1 (12%)	0 (0%)	1 (2%)

Table A4.10. Homework essay: Hierarchical structure.

$N = 55$ essays were collected and analyzed.

C	I	P	W	NR
12 (22%)	2 (4%)	4 (7%)	3 (6%)	34 (62%)

Table A4.11. Precourse survey reanalysis: Relationships.
N = 17 written responses were collected and analyzed.

C	I	P	W	T	NS	NR
3 (18%)	3 (18%)	4 (24%)	0 (0%)	2 (12%)	0 (0%)	5 (29%)

Table A4.12. Precourse survey reanalysis: Relationships.
Questions:
v1: (FIB)
Consider the following:

a. Galaxy
b. Solar System
c. Universe
d. Earth

Rank them by size, from smallest to largest: **D B A C**
v2: (Essay)
Consider the following: galaxy, solar system, universe. Rank them by size, from smallest to largest. Explain your reasoning, including a description of each item.
Answer: solar system < galaxy < universe
A solar system consists of a star, its planets, their moons, and other small debris.
A galaxy contains hundreds of billions of stars (solar systems), gas, dust, and dark matter.
The universe is all of time, space, and its contents. The observable universe includes hundreds of billions of galaxies.

a. Content

	Fall 2008	**Spring 2009 (v2) *N* = 9**	**Spring 2010 (v1) *N* = 17**	**Fall 2010 (v2) *N* = 14**	**Spring 2011 (v2) *N* = 13**	**Total *H* = 1.20 *p* = 0.75 *N* = 53**
C	—	**8 (89%)**	**12 (71%)**	**13 (93%)**	**11 (85%)**	**44 (83%)**
P	—	1 (11%)	4 (24%)	1 (7%)	2 (15%)	8 (15%)
W	—	0 (0%)	1 (6%)	0 (0%)	0 (0%)	1 (2%)

b. Reasoning

	Fall 2008	**Spring 2009 *N* = 9**	**Spring 2010**	**Fall 2010 *N* = 14**	**Spring 2011 *N* = 13**	**Total *H* = 1.33 *p* = 0.51 *N* = 36**
C	—	**5 (56%)**	—	**8 (57%)**	**5 (38%)**	**18 (50%)**
I	—	2 (22%)	—	4 (29%)	5 (38%)	11 (31%)
P	—	2 (22%)	—	1 (7%)	3 (23%)	6 (17%)
W	—	0 (0%)	—	1 (7%)	0 (0%)	1 (3%)
T	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)
NS	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)
NR	—	0 (0%)	—	0 (0%)	0 (0%)	0 (0%)

Table A4.13. Final exam: Solar System and Galaxy (ranking task), reasoning.
Question: Which is bigger, the solar system or the galaxy? Explain your reasoning.
Answer: The galaxy is bigger than the solar system.
A solar system consists of a star, its planets, their moons, and other small debris.
A galaxy contains hundreds of billions of stars (solar systems), gas, dust, and dark matter.

	Fall 2008	**Spring 2009 $N = 9$**	**Spring 2010 $N = 13$**	**Fall 2010 $N = 13$**	**Spring 2011 $N = 13$**	**Total $H = 1.35$ $p = 0.72$ $N = 48$**
C	—	**5 (56%)**	**10 (77%)**	**11 (85%)**	**11 (85%)**	**37 (77%)**
I	—	2 (22%)	0 (0%)	0 (0%)	0 (0%)	2 (4%)
P	—	2 (22%)	2 (15%)	2 (15%)	2 (15%)	8 (17%)
W	—	0 (0%)	1 (8%)	0 (0%)	0 (0%)	1 (2%)
T	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NS	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)
NR	—	0 (0%)	0 (0%)	0 (0%)	0 (0%)	0 (0%)

References

Abazajian, K. *et al*. (194 co-authors) 2009, “The Seventh Data Release of the Sloan Digital Sky Survey,” *ApJS*, 182, 543.

Adams, J. P., Prather, E. E., Slater, T. F., and the CAPER Team. 2005, *Lecture-Tutorials for Introductory Astronomy* (1st ed.), Upper Saddle River, NJ: Prentice-Hall.

American Association for the Advancement of Science [AAAS] 1990, *Science for all Americans*, New York: Oxford University Press.

AAAS 1993, *Benchmarks for Science Literacy*, New York: Oxford University Press.

Bailey, J. M., and Slater, T. F. 2003, “A Review of Astronomy Education Research,” *Astronomy Education Review*, 2, 20.

Bailey, J. M., Coble, K. A., Cochran, G. L., Larrieu, D. M., Sanchez, R., and Cominsky, L. R. 2012, “A Multi-Institutional Investigation of Students’ Preinstructional Ideas About Cosmology,” *Astronomy Education Review*, 11(1), 010302.

Beichner, R. J. 2009, “An Introduction to Physics Education Research,” in *Getting Started in PER*, Reviews in PER Vol. 2, eds. C. Henderson and K. A. Harper, College Park, MD: American Association of Physics Teachers.

Bransford, J. D., Brown, A. L., and Cocking, R. R. (Eds.) 1999, *How People Learn: Brain, Mind, Experience, and School*, Washington, DC: National Academy of Sciences.

Cheek, K. 2012, “Students’ Understanding of Large Numbers as a Key Factor in Their Understanding of Geologic Time,” *International Journal of Science and Mathematics Education*, 10, 1047.

Coble, K., Cominsky, L. R., McLin, K. M., Metevier, A. J., and Bailey, J. M. 2012, “Using the Big Ideas in Cosmology to Teach College Students,” in *Connecting People to Science*, eds. J. B. Jensen, J. G. Manning, M. G. Gibbs, and D. Daou, San Francisco, CA: Astronomical Society of the Pacific, 49.

Coble, K., McLin, K., Bailey, J. M., Metevier, A. J., and Cominsky, L., *The Big Ideas in Cosmology*, Dubuque, IA: Kendall Hunt Publishers/Great River Technology, Inc (in preparation).

Coble, K., Nickerson, M. D., Bailey, J. M., Trouille, L. E., Cochran, G. L., Camarillo, C. T., and Cominsky, L. R., “Investigating Student Ideas about Cosmology II: Composition of the Universe,” *Astronomy Education Review* (submitted).

Cole, S., Percival, W. J., Peacock, J. A., Norberg, P., Baugh, C. M., Frenk, C. S., Baldry, I., Bland-Hawthorn, J., Bridges, T., Cannon, R., Colless, M., Collins, C., Couch, W., Cross, N. J. G., Dalton, G., Eke, V. R., De Propris, R., Driver, S. P., Efstathiou, G., Ellis, R. S., Glazebrook, K., Jackson, C., Jenkins, A., Lahav, O., Lewis, I., Lumsden, S., Maddox, S., Madgwick, D., Peterson, B. A., Sutherland, W., and Taylor, K. 2005, "The 2dF Galaxy Redshift Survey: Power-Spectrum Analysis of the Final Data Set and Cosmological Implications," *MNRAS*, 362, 505.

Delgado, C., Stevens, S. Y., Shin, N., Yunker, M. L., and Krajcik, J. S. 2007, "The Development of Students' Conceptions of Size," paper presented at the Annual Meeting of the National Association for Research in Science Teaching (New Orleans, LA).

Deming, G. L. 2002, "Results from the Astronomy Diagnostic Test National Project," *Astronomy Education Review*, 1(1), 52.

Deming, G. and Hufnagel, B. 2001, "Who's Taking ASTRO 101?," *The Physics Teacher*, 39, 368.

diSessa, A. A. 1988, "Knowledge in Pieces," *in Constructivism in the Computer Age*, eds. G. Forman and P. Pufall, Hillsdale, NJ: Erlbaum, 49.

Donovan, M. S., and Bransford, J. D. (Eds.) 2005, *How Students Learn: Science in the Classroom*, Washington, DC: The National Academies Press.

Duit, R. 2006, *Bibliography – STCSE: Students' and Teachers' Conceptions and Science Education*. Retrieved February 9, 2006, from University of Kiel, Leibniz-Institut für die Pädagogik der Naturwissenschaften, http://www.ipn.uni-kiel.de/aktuell/stcse/.

Fox, M. A. and Hackerman, N. 2003, *Evaluating and Improving Undergraduate Teaching in Science, Technology, Engineering, and Mathematics*, Washington D.C.: The National Academies Press.

Freedman, W. L., Madore, B. F., Gibson, B. K., Ferrarese, L., Kelson, D. D., Sakai, S., Mould, J. R., Kennicutt, R. C., Jr., Ford, H. C., Graham, J. A., Huchra, J. P., Hughes, S. M. G., Illingworth, G. D., Macri, L. M., and Stetson, P. B. 2001, "Final Results from the Hubble Space Telescope Key Project to Measure the Hubble Constant," *ApJ*, 553, 47.

Haynes, M. P., Giovanelli, R., Martin, A. M., Hess, K. M., Saintonge, A., Adams, E. A. K., Hallenbeck, G., Hoffman, G. L., Huang, S., Kent, B. R., Koopmann, R. A., Papastergis, E., Stierwalt, S., Balonek, T. J., Craig, D. W., Higdon, S. J. U., Kornreich, D. A., Miller, J. R., O'Donoghue, A. A., Olowin, R. P., Rosenberg, J. L., Spekkens, K., Troischt, P., and Wilcots, E. M. 2011, "The Arecibo Legacy Fast ALFA Survey: The α.40 H I Source Catalog, Its Characteristics and Their Impact on the Derivation of the H I Mass Function," *AJ*, 142, 170.

Hudgins, D. W., Prather, E. E., Grayson, D. J., and Smits, D. P. 2006, "Effectiveness of Collaborative Ranking Tasks on Student Understanding of Key Astronomy Concepts," *Astronomy Education Review*, 5(1), 1.

Kregenow, J., Rogers, M., and Constas, M. 2010, "Multidimensional Education Research: Managing Multiple Data Streams," *Astronomy Education Review*, 9(1), 010104.

Lelliott, A., and Rollnick, M. 2010, "Big Ideas: A Review of Astronomy Education Research 1974–2008," *International Journal of Science Education*, 32, 1771.

Miller, B. W., and Brewer, W. F. 2010, "Misconceptions of Astronomical Distances," *International Journal of Science Education*, 32, 1549.

National Research Council [NRC] 1996, *National Science Education Standards*, Washington, DC: National Academy Press.

NRC 2003, *Learning and Understanding: Improving Advanced Study of Mathematics and Science in U.S. High Schools*, Washington, DC: National Academy Press.

NRC 2012, *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*, Washington, DC: The National Academies Press.

Nolan, P. L. *et al*. (67 co-authors) 2012, “Fermi Large Area Telescope Second Source Catalog,” *ApJS*, 199, 31.

The Office of Institutional Effectiveness and Research, 2011, *Chicago State University 2011 Electronic Fact Book*, Chicago, IL: Chicago State University. Available online at http://www.csu.edu/IER/documents/FactBook_2011-12.pdf

O’Kuma, T. L., Maloney, D. P., and Hieggelke, C. J. 2000, *Ranking Task Exercises in Physics*, Upper Saddle River, NJ: Prentice-Hall.

Palen, S., n.d., *Distance Measures* [lab, part 3], University of Washington, Dept. of Astronomy, http://www.astro.washington.edu/courses/labs/clearinghouse/labs/Distance/distance.html

Pasachoff, J. M. 2002, “What Should College Students Learn? Phases and Seasons? Is Less More or Is Less Less?,” *Astronomy Education Review*, 1(1), 124.

Prather, E. E., Slater, T. F., and Offerdahl, E. G. 2002, “Hints of a Fundamental Misconception in Cosmology,” *Astronomy Education Review*, 1(2), 28.

Rubin, H. J., and Rubin, I. S. 2005, *Qualitative Interviewing: The Art of Hearing Data*, 2nd ed., Thousand Oaks, CA: Sage.

Rudolph, A. L., Prather, E. E., Brissenden, G., Consiglio, D., and Gonzaga, V. 2010, “A National Study Assessing the Teaching and Learning of Introductory Astronomy Part II: The Connection between Student Demographics and Learning,” *Astronomy Education Review*, 9, 010107.

Sayre, E. C., and Heckler, A. F. 2009, “Peaks and Decays of Student Knowledge in an Introductory E & M Course,” *Physical Review Special Topics - Physics Education Research*, 5, 013101.

Skrutskie, M. F., Cutri, R. M., Stiening, R., Weinberg, M. D., Schneider, S., Carpenter, J. M., Beichman, C., Capps, R., Chester, T., Elias, J., Huchra, J., Liebert, J., Lonsdale, C., Monet, D. G., Price, S., Seitzer, P., Jarrett, T., Kirkpatrick, J. D., Gizis, J. E., Howard, E., Evans, T., Fowler, J., Fullmer, L., Hurt, R., Light, R., Kopan, E. L., Marsh, K. A., McCallon, H. L., Tam, R., Van Dyk, S., and Wheelock, S. 2006, “The Two Micron All Sky Survey (2MASS),” *AJ*, 131, 1163.

Slater, T. F., Adams, J. P., Brissenden, G., and Duncan, D. 2001, “What Topics Are Taught in Introductory Astronomy Courses?,” *The Physics Teacher*, 39, 52.

Smithsonian Astrophysical Observatory 2003, *Cosmic Survey: What Are Your Ideas About the Universe?*, http://www.cfa.harvard.edu/seuforum/download/CosmicSurvey2003.pdf

Spencer, B. H., and Guillaume, A. M. 2006, “Integrating Curriculum through the Learning Cycle: Content-Based Reading and Vocabulary Instruction,” *Reading Teacher*, 60, 206.

Tretter, T. R., Jones, M. G., and Minogue, J. 2006, “Accuracy of Scale Conceptions in Science: Mental Maneuverings across Many Orders of Spatial Magnitude,” *Journal of Research in Science Teaching*, 43, 1061.

Trouille, L. E., Coble, K., Cochran, G. L., Bailey, J. M., Camarillo, C. T., Nickerson, M. D., and Cominsky, L. R. “Investigating Student Ideas about Cosmology III: Big Bang Theory, Expansion, Age, and History of the Universe,” *Astronomy Education Review* (submitted).

Wallace, C. S., Prather, E. E., and Duncan, D. K. 2012, “A Study of General Education Astronomy Students’ Understandings of Cosmology. Part IV. Common Difficulties Students Experience with Cosmology,” *Astronomy Education Review*, 11, 010104.

ÆR
010102-1–010102-54