An investigation into introductory astronomy students' difficulties with cosmology, and the development, validation, and efficacy of a new suite of cosmology lecture-tutorials

by

Colin S. Wallace

M.S., University of Colorado at Boulder, 2007

B.S., Case Western Reserve University, 2005

A thesis submitted to the
Faculty of the Graduate School of the
University of Colorado in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
Department of Astrophysical and Planetary Sciences
2011

This thesis entitled:
An investigation into introductory astronomy students' difficulties with cosmology, and the development, validation, and efficacy of a new suite of cosmology lecture-tutorials
written by Colin S. Wallace
has been approved for the Department of Astrophysical and Planetary Sciences

Prof. Edward E. Prather

Dr. Doug Duncan

Prof. Noah Finkelstein

Prof. Nicholas Schneider

Prof. Fran Bagenal

Date

The final copy of this thesis has been examined by the signatories, and we find that both the content and the form meet acceptable presentation standards of scholarly work in the above mentioned discipline.

IRB protocol #1208.10

Wallace, Colin S. (Ph.D., Astrophysical and Planetary Sciences)

An investigation into introductory astronomy students' difficulties with cosmology, and the development, validation, and efficacy of a new suite of cosmology lecture-tutorials

Thesis directed by Prof. Edward E. Prather

This study reports the results of the first systematic investigation into Astro 101 students' conceptual and reasoning difficulties with cosmology. We developed four surveys with which we measured students' conceptual knowledge of the Big Bang, the expansion and evolution of the universe, and the evidence for dark matter. Our classical test theory and item response theory analyses of over 2300 students' pre- and post-instruction responses, combined with daily classroom observations, videotapes of students working in class, and one-on-one semi-structured think-aloud interviews with nineteen Astro 101 students, revealed several common learning difficulties. In order to help students overcome these difficulties, we used our results to inform the development of a new suite of cosmology lecture-tutorials. In our initial testing of the new lecture-tutorials at the University of Colorado at Boulder and the University of Arizona, we found many cases in which students who used the lecture-tutorials achieved higher learning gains (as measured by our surveys) at statistically significant levels than students who did not. Subsequent use of the lecture-tutorials at a variety of colleges and universities across the United States produced a wide range of learning gains, suggesting that instructors' pedagogical practices and implementations of the lecture-tutorials significantly affect whether or not students achieve high learning gains.

Dedication

This dissertation is dedicated to my parents and my wife. My parents, Jim and Sandy, inspired and nurtured my love of science, even though they are not scientists themselves. My wife, Alissa, has been my supportive and loving companion throughout my time in graduate school, despite the fact that we have had to postpone many life experiences due to my quixotic quest for this degree. To them, I extend my love, appreciation, and thanks.

Acknowledgements

There are many people who contributed to this study. Edward Prather, Doug Duncan, and Noah Finkelstein have provided invaluable guidance and advice throughout this project. Furthermore, Edward Prather and I are equal coauthors of the cosmology lecture-tutorials. This project has also benefited from my interactions with John Bally, Derek Briggs, Gina Brissenden, Timothy Chambers, Sebastien Cormier, Andrew Hamilton, Seth Hornstein, Steve Pollock, the Physics Education Research Group at the University of Colorado, and the Center for Astronomy Education. Without these people, this dissertation would only be a shadow of its current form. I frequently use the pronoun "we" throughout this dissertation as an acknowledgment of the contributions of all these people.

Of course, I must also thank the dozens of instructors and thousands of students who participated in this study. Clearly, their participation was essential to the success of this project.

I also owe a special thanks to Andrew Hamilton and John Stocke. Although they played limited roles in this project, they helped guide me to the point of candidacy during a time in which I frequently doubted my ability to earn a Ph.D. Without their support, I would not be where I am today.

This material is based upon work supported by the National Science Foundation under Grants No. 0833364 and 0715517, a CCLI Phase III Grant for the Collaboration of Astronomy Teaching Scholars (CATS). Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author and do not necessarily reflect the views of the National Science Foundation.

Contents

Chapter

Appendix

Tables

Table

Figures

Figure

Chapter 1

Introduction

Modern cosmology is one of the most dynamic fields of astrophysics. Over the past century, cosmologists have addressed topics including the beginning of the universe, the composition of the universe, and the evolution and eventual fate of the universe. Such topics were once confined to the domains of philosophy and theology, but theoretical and technological advances have allowed these issues to be tackled from a scientific perspective. By the end of the twentieth century and the beginning of the twenty-first century, we had multiple lines of evidence all pointing toward a concordant cosmological model (Tegmark, Zaldarriaga, and Hamilton 2001). This model permits cosmologists to make detailed descriptions of the components, history, and evolution of our universe – leading some to declare that we are in an era of "precision cosmology" (Primack 2005).

Yet many people find many aspects of cosmology esoteric. For some, the recondite nature of cosmology is due to the heavy mathematical burdens of general relativity and quantum mechanics, both of which play important roles in our descriptions of the universe. Others frequently misunderstand the very concepts of cosmology, including many professional astronomers. Some cosmologists, such as Davis and Lineweaver (2004), have written technical clarifications of several points of confusion for their colleagues. Such difficulties among professional astronomers beg the question of what difficulties our students experience when we try to teach them cosmology.

1.1 Astronomy Education Research

Answering this question falls into the domain of astronomy education research (AER). Following the well-blazed path of physics education research (PER), the burgeoning field of AER traditionally focuses on two areas of research. First, astronomy education researchers attempt to uncover and model student difficulties on common astronomy topics. Such topics include lunar phases, seasons, the motions of planets, and gravity, among others (see Bailey and Slater 2005 for a relatively recent summary of AER research). Second, they also develop and study the tools, techniques, and implementation practices that affect student learning. Several research-validated instructional strategies have emerged from this effort, such as lecture-tutorials (Prather *et al.* 2008) and ranking tasks (Hudgins *et al.* 2006). These endeavors typically focus on general education, college-level introductory astronomy courses taken primarily by non-science majors, hereafter referred to as Astro 101.

Why has Astro 101 been the predominant focus of AER? One reason is because such courses reach a wide variety of students. Up to a quarter of a million students take an Astro 101 course in the United States each year (Fraknoi 2002). Demographic studies reveal that Astro 101 students are broadly representative of the American undergraduate population (Deming and Hufnagel 2001; Rudolph *et al.* 2010). As others have noted (Prather *et al.* 2009; Zeilik *et al.* 1997), Astro 101 is the terminal science course in life for many of these students. It thus represents an important and final opportunity for these students – who will become the nation's future politicians, journalists, and business leaders – to develop scientific literacy. Furthermore, up to a quarter of Astro 101 students are education majors (Rudolph *et al.* 2010). Since these students will eventually be hired to teach science to future generations, Astro 101 instructors have a responsibility to model effective pedagogical techniques, in addition to improving these students' science content knowledge. These are the reasons why much of AER is devoted to Astro 101.

Yet even in the well-studied area of Astro 101, AER has very little to say about students' difficulties in cosmology. There are only a handful of studies that, in whole or in part, address student difficulties in cosmology (Comins 2001; Lightman and Miller 1989; Lightman, Miller, and Leadbeater 1987; Prather, Slater, and Offerdahl 2002; Simonelli and Pilachowski 2003). This is despite the fact that cosmology is one of the most commonly taught topics in Astro 101 (Slater *et al.* 2001). This study is an attempt to fill this hole in the AER literature.

1.2 Overview of Dissertation

This dissertation is one of the first systematic studies of Astro 101 students' difficulties in cosmology (Bailey *et al.* 2011 and Coble *et al.* 2011 describe another study complementary to our's). As such, we have two major questions we want to answer.

First, what are the common conceptual and reasoning difficulties Astro 101 students encounter when they study cosmology? To answer this question, we had to design and analyze a new set of surveys on conceptual cosmology knowledge. This was necessary because there are currently no research-validated instruments suitable for this purpose. We followed Wilson's (2005) four-step procedure for survey design and interpretation. This process is described and exemplified throughout this dissertation.

However, this dissertation focuses on more than just identifying common student difficulties. We also created new resources Astro 101 instructors can use in their classrooms to help students overcome their difficulties. Specifically, we created a new suite of five lecture-tutorials, all of which focus on some aspect of cosmology. These lecture-tutorials are designed to be like the original thirty-eight lecture-tutorials (only two of which address cosmology) in the book *Lecture-Tutorials for Introductory Astronomy* (Prather *et al.* 2008). What is a lecture-tutorial? A lecture-tutorial is a two to six page worksheet comprised of Socratic-style questions on a topic research has shown students struggle with (Prather *et al.* 2005). Lecture-tutorials are designed to be integrated into the lecture portion of a class (Prather *et al.* 2005). Research on the original lecture-tutorials shows they produce larger learning gains than lecture alone (Prather *et al.* 2005; LoPresto and Murrell 2009). Thus, the second question we seek to answer is "Do students who use the new cosmology lecture-tutorials achieve larger learning gains than students who do not?"

This dissertation is organized as follows. Chapter 2 summarizes current perspectives on the nature of learning. This chapter is fundamental for understanding why a particular resource, such as the lecture-tutorials, might produce larger learning gains than more traditional forms of instruction. It also outlines theoretically-supported and research-tested methods for helping students learn. These methods played important roles in our design of the cosmology lecture-tutorials. Chapter 3 describes and defends our selection of cosmology topics for this study: the expansion and evolution of the universe, the Big Bang, and evidence for dark matter. Chapter 3 also contains a brief review of the cosmology relevant to these topics. Chapters 4 and 5 both focus on the design processes pertinent to this study: Chapter 4 describes the process we used to design our surveys of conceptual cosmology knowledge, while Chapter 5 describes the design of the cosmology lecture-tutorials. Chapters 6-8 constitute the heart of this dissertation. They describe our analyses of students' survey responses and present the survey data necessary to answer the two main questions of this study. These chapters also describe other forms of data we collected, such as one-on-one student interviews and classroom observations, to supplement the survey data. Each of these chapters focuses on the data we collected for a particular semester: Chapter 6 presents data from the fall 2009, Chapter 7 presents data from the spring 2010, and Chapter 8 presents data from the fall 2010. Chapters 6 and 7 further discuss the revisions we made to our surveys and lecture-tutorials after the fall 2009 and spring 2010, respectively, in responses to potential problems revealed by the data. Chapter 9 summarizes the results of this study and presents our conclusions.

Chapter 2

Theories of Learning and Conceptual Change

Reams of AER and PER studies are devoted to elucidating common student difficulties on a plethora of topics. Such common difficulties are often explicitly addressed by research-based curricular interventions, such as the original *Lecture-Tutorials for Introductory Astronomy* (Prather *et al.* 2008). These curricular interventions often produce larger learning gains than lecture, by itself, can accomplish.

For example, Prather *et al.* (2005) measured the efficacy of the *Lecture-Tutorials* via students' responses to sixty-eight multiple choice questions. These questions all possessed attractive distractors and emphasized conceptual understanding over factual recall. They were administered to students at the beginning of the course and then a subset of these questions were selected based on a given day's topic and administered post-lecture and post-lecture-tutorial. The average pre-instruction score was 30% correct. After lecture, this average improved to 52% and, post lecture-tutorial, it increased to 72% (Prather *et al.* 2005). A subsequent study at a different institution supports Prather *et al.*'s (2005) claim that the *Lecture-Tutorials* lead to higher learning gains than traditional lectures (LoPresto and Murrell 2009).

Why do we need to research common student difficulties in astronomy? Furthermore, why should research-based interventions, such as the original *Lecture-Tutorials*, be more effective than lecture alone? I address these question in this chapter by reviewing the extensive and ever-growing literature on theories of learning and conceptual change. I begin in Section 2.1 by showing that the *Lecture-Tutorials* and their attendant research are part of a long line of studies arguing for increasing the *interactive engagement* of students in class. I then discuss how the success of interactive engagement strategies may be understood from two complementary perspectives on learning: the cognitive perspective (Section 2.2) and the situated perspective (Section 2.3). This chapter concludes with my summary of what these theories of learning have to say about the effectiveness of interactive engagement (generally) and the *Lecture-Tutorials* (specifically).

2.1 The Importance of Interactive Engagement

For years, scientists gathered evidence that traditional styles of instruction, especially lecture, were poor methods for promoting student learning. For example, the PER group at the University of Washington has spent decades cataloging common student difficulties in all areas of physics, such as kinematics (Trowbridge and McDermott 1980; Trowbridge and McDermott 1981), graph interpretation (McDermott, Rosenquist, and van Zee 1987), buoyancy (Heron 2004a), electric circuits (McDermott and Shaffer 1992a), and special relativity (Scherr, Shaffer, and Vokos 2001; Scherr, Shaffer, and Vokos 2002), among many others (see McDermott and Redish 1999 for more). Many of the conceptual and reasoning difficulties uncovered by the Washington group persist post-lecture unless students spend class time engaged in activities that specifically target these difficulties (Heron 2004b; Scherr, Shaffer, and Vokos 2002; Shaffer and McDermott 1992b). These results cannot be dismissed by saying that students simply need more polished lectures; while students may appreciate the organization of a good lecture, similar post-lecture struggles are reported from multiple institutions across multiple years. In one case, Clement (1982) reports that, post-lecture, many introductory physics students continue to reason that there is always a force in the direction of motion. This result held even for students of a highly regarded teacher and for high-achieving students who subsequently progressed into advanced science, engineering, and mathematics courses (Clement 1982). Furthermore, these fundamental conceptual problems often persist among students who are adroit problem-solvers. Mazur (1997) warns that problem-solving prowess is, in many cases, not a sufficient condition for conceptual mastery.

These results are not limited to physics. AER studies likewise point to the ineffectiveness of traditional instruction. For example, Lightman and Sadler (1993) looked at the pre- and post-course responses of 330 eighth to twelfth grade students to a sixteen item multiple choice astronomy test. The sixteen items covered topics such as motions of the Sun and Moon, light, and gravity, and included distractors that reflect common errors in students' thinking. These questions were also divided between those that require factual recall and those that require conceptual understanding. Lightman and Sadler found the average pre- to post-course improvement (measured by subtracting the pre-course from the post-course score) is negligible. Furthermore, they found that while the average post- minus pre-course difference was positive for the factual questions, it was negative for the conceptual questions. These problems do not appear to result from teachers overestimating the pre-instruction knowledge and abilities of their students; the teachers of the students in Lightman and Sadler's study did a good job, on average, of predicting their students' average pre-course performances on each item. However, these teachers grossly overestimated their students' post-course performances, especially on conceptual questions (Lightman and Sadler 1993). Lightman and Sadler note that traditional forms of instruction often fail to overcome deeply held ideas that conflict or interfere with the material presented by the instructor.

Many physicists and astronomers became acutely aware of the problems with traditional instruction when, in 1998, the *American Journal of Physics* published a landmark paper by Richard Hake. Hake (1998a) looked at the pre- and post-instruction class averages on the Mechanics Diagnostic Test (Halloun and Hestenes 1985), the Force Concept Inventory (Hestenes, Wells, and Swackhamer 1992), and the Mechanics Baseline Test (Hestenes and Wells 1992) for 6542 students in sixty-two introductory physics classes. His sample included both calculus and algebra-based classes and classes taught at high schools, colleges, and universities across the United States. To measure the success of each class, Hake used the normalized gain statistic, which is defined as

$$<g>=\frac{S_f - S_o}{100\% - S_o} \ , \tag{2.1}$$

where S_f is the average post-instruction score for a class and S_o is the average pre-instruction score. This statistic represents the average amount by which a class improved as a fraction of the total amount by which it could have improved (Hake 1998a; Prather *et al.* 2009).

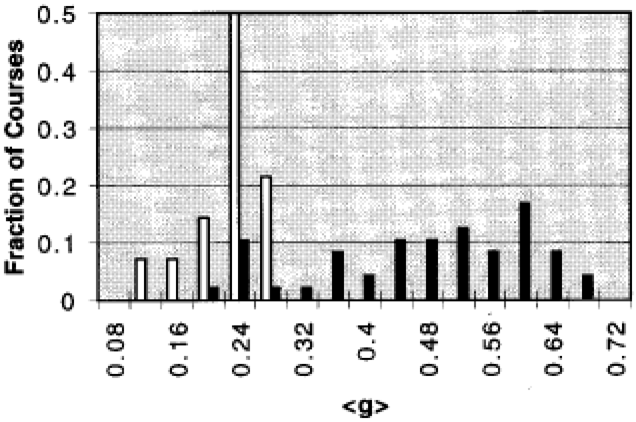


Figure 2.1: Hake's (1998a) histogram of average normalized gains ($< g >$) for traditional classes (white bars) and interactive engagement classes (black bars).

Hake's crucial result came when he compared the normalized gains of traditional and interactive engagement classes. He defined a traditional class as one that primarily and passively transmits information via lecture, cookbook labs, and problem-solving exams (Hake 1998a). He defines interactive classes as those which used in-class activities to promote conceptual understanding via student-student and student-teacher discussion and feedback (Hake 1998a). His results are reproduced in Figure 2.1. The fourteen traditional classes in his study have an average normalized learning gain of 0.23 ± 0.04 while the forty-eight interactive engagement classes have an average normalized learning gain of 0.48 ± 0.14 (Hake 1998a). The limitations of traditional lecture are underscored by the relatively small spread in the data for the traditional courses. Hake notes that the larger spread in the distribution of learning gains for interactive courses may be explained, in part, by how effectively instructors implement interactive engagement strategies. However, as Figure 2.1 makes evident, even the very worst interactive engagement classes achieve learning gains comparable to the very best traditional classes.

This study has been repeated in the context of introductory astronomy. Prather *et al.* (2009) looked at the pre- and post-instruction responses of approximately 4000 introductory astronomy students to the Light and Spectroscopy Concept Inventory (Bardar *et al.* 2007). These students were drawn from classes taught by thirty-six instructors representing thirty-one institutions, including two and four year colleges as well as universities, across the United States (plus one from Ireland). The surveyed classes ranged in size from fewer than twenty-five students to 100 or more students. Prather *et al.* (2009) found that, regardless of institution or class-size, all courses in which the instructor spent less than a quarter of class time on interactive activities (self-reported) failed to achieve average normalized learning gains greater than 0.30 (see Figure 2.2).

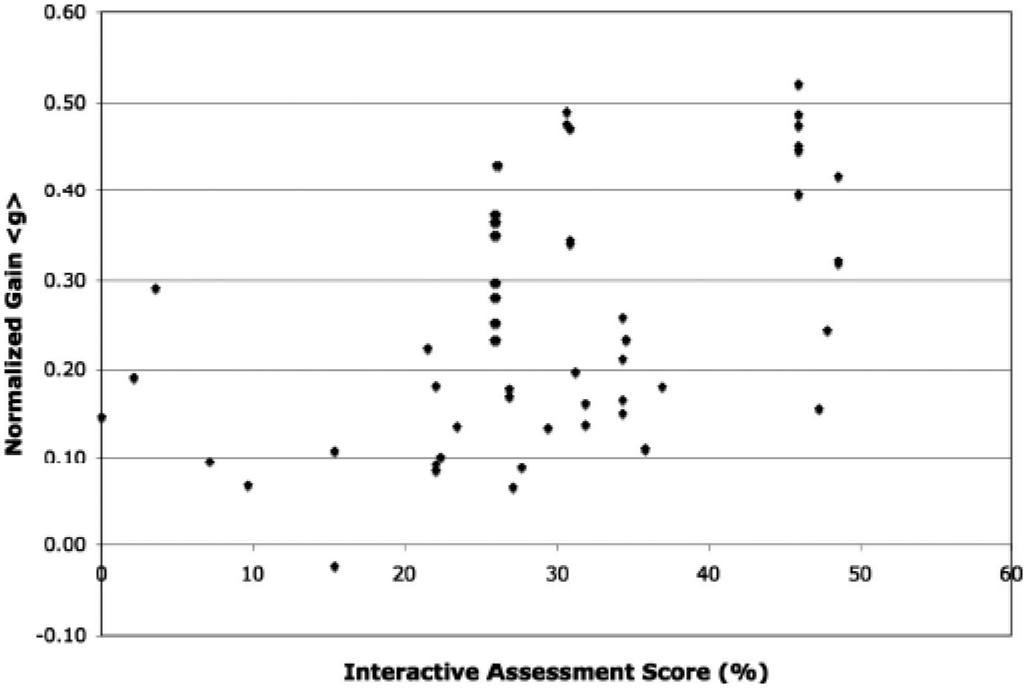


Figure 2.2: Prather *et al.*'s (2009) plot of average normalized gains ($< g >$) as a function of (instructor reported) class time devoted to interactive engagement.

Researchers in PER and AER have developed many research-validated strategies and resources for promoting interactive engagement and improved student learning. Many focus on students' conceptual knowledge. For example, think-pair-share (Lyman 1981), also known as peer instruction (Mazur 1997), is a process by which students are given a question, asked to vote individually on the answer, and, if between 50-80% are correct, told to defend their answer to their neighbor before revoting (see also Duncan 2007; Green 2003; and Prather and Brissenden 2009). Tutorials are worksheet-based activities containing sequences of conceptual questions. These sequences guide student discussions in an effort to help them overcome common conceptual difficulties and construct their own understandings of a topic (Heron 2004b; McDermott and Shaffer 1992b; Prather *et al.* 2005; Scherr, Shaffer, and Vokos 2002). Interactive lecture demonstrations help make in-class demonstrations more than just a passive activity for students (Sokoloff and Thornton 1997). Instructors who want to promote problem-solving skills and quantitative reasoning may adopt activities such as context rich problems (Heller, Keith, and Anderson 1992; Heller and Hollabaugh 1992) or ranking tasks (Hudgins *et al.* 2006). These are just a subset of the interactive engagement strategies available to physics and astronomy instructors (for more details, see Redish 2003 and references therein).

These interactive engagement activities appear to benefit all students, regardless of their backgrounds. For example, a recent study by Rudolph *et al.* (2010) used a multiple regression analysis to explore the relationship between interactive instruction, student learning, and a host of demographic characteristics. They found that interactive engagement leads to improved learning for both genders, across ethnicities, and irrespective of one's primary language or academic background.

In summary, decades of AER and PER studies show that common student difficulties are frequently unaffected by traditional lecture-based instruction. Some sort of interactive engagement is required to effect deeper levels of learning. These results are not limited to a particular class, race, or gender: All learn and benefit from interactive lessons. These results say something profound about how people learn (NRC 2000).

In the following sections, I discuss two models of how people learn: the cognitive (Section 2.2) and the situated (Section 2.3) models. These models are necessary for three reasons. First, they help explain why interactive engagement improves student learning. Second, they illuminate the processes by which students learn. Finally, the help identify promising activities and strategies for promoting increased learning.

2.2 Cognitive Models of Learning

2.2.1 Overview and Key Results

Cognitive models of learning arose in reaction against the behaviorist view of learning. Behaviorism dominated psychology during the first part of the twentieth century. Its adherents were concerned that hypotheses based on conjectured internal mental states are ultimately too subjective and metaphysical (Anderson, Reder, and Simon 1998; NRC 2000). They defined learning as the process by which observed responses are associated with external stimuli (Anderson, Reder, and Simon 1998; NRC 2000; Shepard 2000). While this approach had some success, it eventually waned in popularity as researchers realized that some behaviors can only be understood in terms of unobservable mental states and processes (Anderson, Reder, and Simon 1998; NRC 2000; Redish 1994).

What followed behaviorism has become known as the "cognitive revolution" (Anderson, Reder, and Simon 1998; Redish 1994; Shepard 2000). Cognitivists draw on the findings and methods of a number of disciplines, such as neuroscience, anthropology, linguistics, and psychology, in order to develop and test their ideas on human thinking and learning (NRC 2000). They are primarily concerned with how people organize information into mental structures and how those mental structures subsequently influence the learning and processing of new information (Anderson, Reder, and Simon 1998; NRC 2000; Hammer *et al.* 2005; Redish 1994).

One of the primary findings of cognitive research is that students do not enter the classroom as *tabula rasa*; the knowledge, intuitions, and beliefs they bring with them exert complex and profound influences over what they learn (Carey 1988; diSessa 1993; Elby 2001; Minstrell 1992; NRC 2000; Posner *et al.* 1982; Redish 1994; Vosniadou 1994). For example, Vosniadou (1994) describes how children reconcile their intuition that the Earth is flat with teachings claiming the Earth is round. While some children (especially older children) adopt the scientifically correct model of a spherical Earth, others fuse their intuitions with what they are taught to form non-expert-like synthetic models. These synthetic models include ideas such as the Earth is a flat disk, the Earth is a hollow sphere (in which we live on a plane in its interior), and the Earth is both a flat plane on which we live as well as a separate, spherical planet that exists in the sky (Vosniadou 1994).

A second major finding of cognitive research is that failing to account for students' prior knowledge frequently results in ephemeral and superficial learning (McDermott 1991; NRC 2000; Redish 1994). Vosniadou (1994) highlights three different "failures in learning": inconsistencies, inert knowledge, and misconceptions. Inconsistencies occur when students simply add conflicting pieces of information to their knowledge base. Inert knowledge occurs when information that conflicts with their prior knowledge is only activated in a limited number of circumstances, such as a final exam. Misconceptions arise when students merge their prior knowledge with what they are learning to produce non-expert-like models, such as the various Earth models mentioned above. The National Research Council report *How People Learn* summarizes the situation as follows:

> "Students come to the classroom with preconceptions about how the world works. If their initial understanding is not engaged, they may fail to grasp the new concepts and information that are taught, or they may learn them for the purposes of a test but revert to their preconceptions outside the classroom" (NRC 2000, pp.14-5).

The knowledge, intuitions, and beliefs that play central roles in how students organize and interpret information are often especially impervious to change (Redish 1994; Strike and Posner 1992; Vosniadou 1994).

However, students' learning struggles may be due to more than just possessing non-expert-like ideas. As many others have noted, how students organize and activate their knowledge is just as important (diSessa 1993; NRC 2000; Redish 1994). Experts typically organize their knowledge around major principles or "big ideas" and understand the contexts in which certain pieces of knowledge are applicable (NRC 2000). Novices, in contrast, tend to lack a sense of which ideas are central and which are secondary to a discipline, and how all the various ideas relate to one another (NRC 2000). They also tend to focus on the surface features of a particular situation when deciding which bit of information to use (NRC 2000). Redish (1994) makes the following analogy for how physics novices organize their knowledge:

> "It is as if physics were a collection of equations on fallen leaves. One might hold $s = 1/2gt^2$, another $F = ma$, and a third $F = -kx$. These are each considered of equivalent weight, importance, and structure. The only thing one needs to do when solving a problem is to flip through one's collection of leaves until one finds an appropriate equation" (p. 799).

From this perspective, effective learning requires instruction that not only accounts for students' ideas, but also addresses how to organize those ideas.

Given that students' prior knowledge plays a central role in what and how they learn, how can instructors use or change what is already inside students' heads? This is the focus of the next two subsections, which deal with cognitive models of conceptual change (Section 2.2.2) and instructional strategies grounded in cognitive research (Section 2.2.3).

2.2.2 Cognitive Models of Concepts and Conceptual Change

Conceptual change is a term used by multiple researchers (e.g. Carey 1988; diSessa and Sherin 1998; and Posner *et al.* 1982), although its meaning is not unambiguous, as I discuss below. For now, I will provisionally accept Strike and Posner's (1992) description of a concept as mental structure that plays "a generative or organizing role in thought" (p. 148). Conceptual change refers to the process by which experiences and information alter these concepts. Conceptual change is different from the simple amalgamation of new information onto existing concepts (Posner *et al.* 1982; Vosniadou 1994).

Some have likened the process of conceptual change in individuals to the process of Kuhnian scientific revolutions occasionally experienced by scientific communities (Carey 1988; Posner *et al.* 1982; Vosniadou 1994). In a seminal paper, Posner *et al.* (1982) transform this analogy into a fullblown model of conceptual change. They start from the assumption that people "comprehend and accept ideas because they are seen as intelligible and rational" (Posner *et al.* 1982, p. 212). Posner *et al.* (1982) note that while a person can sometimes understand new information using their existing concepts (what Posner *et al.* call *assimilation*), sometimes new information causes a radical restructuring or reorganization of a person's concepts (a process they call *accommodation*).

In order for accommodation to occur, the following conditions must be met (Posner *et al.* 1982; Strike and Posner 1985; Strike and Posner 1992):

1. The person must be dissatisfied with her current concepts;
2. The new concept must be intelligible;
3. The new concept must be plausible (at least initially);
4. The new concept must be fruitful; that is, it must suggest new forms of inquiry or ways of looking at the world.

The authors are careful to note that accommodation is not necessarily an instantaneous event and, in fact, a person may require passing through several iterations of unsuccessful assimilation before she is prepared to accommodate a new concept (Posner *et al.* 1982; Strike and Posner 1985).

In its original formulation, this view of conceptual change views concepts as robust mental structures that are the fundamental units of cognition and are accepted, manipulated, and discarded as a whole (Strike and Posner 1992). This is consistent with much of the "misconceptions" literature that was published around the same time. For example, multiple studies by McCloskey and his colleagues examined students' naïve ideas of motion (McCloskey, Caramazza, and Green 1980; McCloskey, Washburn, and Felch 1983). They interpreted many students' incorrect statements as evidence that they possessed concepts of motion similar to the medieval impetus theory (McCloskey, Caramazza, and Green 1980; McCloskey, Washburn, and Felch 1983). Clement (1982) likewise traced students' difficulties in introductory mechanics to their possession of pre-Newtonian concepts. Yet subsequent work questions this view of robust misconceptions.

Instead of robust concepts, several researchers suggest that students' views of the world are shaped by the in-the-moment, context-dependent activiation (and supression) of a suite of cognitive elements (diSessa 1993; Hammer *et al.* 2005; Minstrell 1992). One of the most famous examples of these models is diSessa's (1993) idea of *phenomenological primatives* (p-prims). diSessa (1993) claims that individuals possess a "sense of mechanism" that allows them to make rapid judgements about what sort of events they expect to happen in the physical world. For example, our sense of mechanism tells us that a ball rolling along the floor will eventually slow and come to a stop. Rather than constituting any sort of formal theory (such as the impetus theory), diSessa claims that our sense of mechanism is comprised of many small, weakly organized elements (hence primatives) that arise from abstracting everyday events (hence phenomenological). A single p-prim is neither correct nor incorrect, although it may be activated at inappropriate junctures. For example, one of diSessa's hypothetical p-prims, which he calls Ohm's p-prim, may be expressed as "an agent...that acts against a resistance to produce some sort of result" (diSessa 1993, p. 126). This can be an appropriate p-prim to activate when thinking about Newton's second law, where the force is the "agent," the mass is the "resistance," and the acceleration is the "result." However, the p-prim becomes inappropriate when acceleration is replaced by velocity as the "result." From this perspective, learning involves improving one's understanding about when to activate and when to supress certain p-prims, as well as how to organize them such that one's intuitions are subsumed into a few deep physical laws (diSessa 1993).

diSessa's p-prims share a number of similarities with what Minstrell (1992) calls "facets" and what Hammer *et al.* (2005) call "resources." A facet is a piece of knowledge, strategy, or reasoning approach that a student applies in various situations (Minstrell 1992). One example facet highlighted by Minstrell is "more...means more..." as in "The more bulbs, the more resistance in the circuit" (Minstrell 1992, p. 112). A resource is what Hammer *et al.* (2005) consider to be a fundamental, fine-grained unit of cognition. What is traditionally referred to as a concept may instead be the momentary activation of multiple resources in response to a given situation (Hammer *et al.* 2005). Since "resources" is meant to be a more general term than p-prims, facets, and similar cognitive models (Frank, Kanim, and Gomez 2008; Hammer *et al.* 2005; Heron 2004a), and since a detailed comparison of these models is beyond the scope of this overview, I will hereafter refer to *cognitive resources* when speaking of fine-grained models of cognition.

Multiple studies provide evidence in favor of the resources perspective. For example, Elby (2000) hypothesizes a resource he christens "what-you-see-is-what-you-get" (WYSIWYG). WYSI-

WYG may be stated as "x means x", where x is some feature or attribute of interest (e.g. a student who activates WYSIWYG may interpret peaks and valleys on graphs as representing peaks and valleys on a terrain, even if the graph plots velocity versus time; Elby 2000). Elby suggests that this resource tends to be activated whenever there is a compelling visual attribute – that is, whenever there is a particular feature (or features) that immediately captures one's attention. Elby goes on to claim that the activation of this resource is a better explanation than the possession of a robust misconception for the pattern of students' responses to several questions. In one case, Elby discusses students responses to the question shown in Figure 2.3. He says that if students possess a misconception that causes them to read velocity graphs as if they were position graphs, then they should answer both questions incorrectly. On the other hand, if students do not possess this misconception but instead activate the WYSIWYG resource when they notice the compelling visual attribute of the lines' intersection, then they should incorrectly answer the second question more often than the first (Elby 2000). In the pilot study, one student answered both questions correctly, two answered both incorrectly, and six answered the first question correctly and the second incorrectly (Elby 2000). Elby argues that this pilot study is consistent with his hypothesis about the WYSIWYG resource (Elby 2000).

In another case, Frank, Kanim, and Gomez (2008) compare the responses two different phrasings of the same question evoke from students. For example, the physical situations in Figures 2.4 and 2.5 are identical. The only difference between these two questions is whether the experiment is described in terms of velocity or distance traveled. Frank, Kanim, and Gomez (2008) argue that if students possess robust misconceptions, such as an impetus theory of motion, then the distribution of their responses should be the same regardless of the experiment's description. Instead, they found that the wording of the question measurably changes the distribution of responses at a statistically significant level (Frank, Kanim, and Gomez 2008). When the question is phrased using velocity, students are more likely to answer that the ball in experiment 3 takes the least amount of time to hit the floor. When distance is used instead of velocity, students are more likely to argue that the ball in experiment 3 takes the longest amount of time to hit the floor. Frank, Kanim,

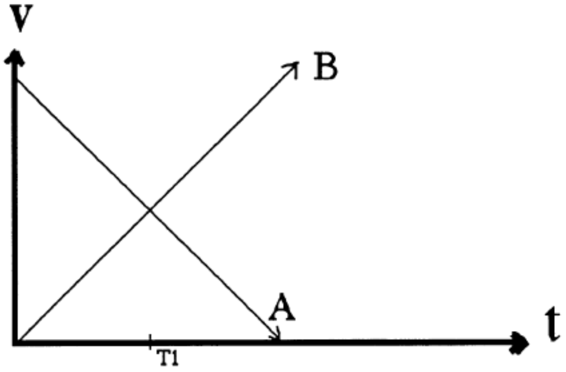


Figure 2.3: Cars A and B start at the same position and move according to the graph of speed versus time shown above.
(1) Is car A going forward or backward? What about car B?
(2) What happens at time $T1$? Circle the correct response.
(a) Car B is ahead.
(b) Car A is ahead.
(c) Neither car is ahead; car B and car A cross each other.
This question is taken from Elby (2000).

and Gomez interpret these results as evidence that the two alternate wordings of the question cue different resources: The velocity wording preferentially cues the resource "going faster means taking less time," while the distance wording preferentially cues the resource "going farther means taking more time." Frank, Kanim, and Gomez (2008) report additional experiments they conducted that support a resource model of cognition for students instead of a misconceptions model.

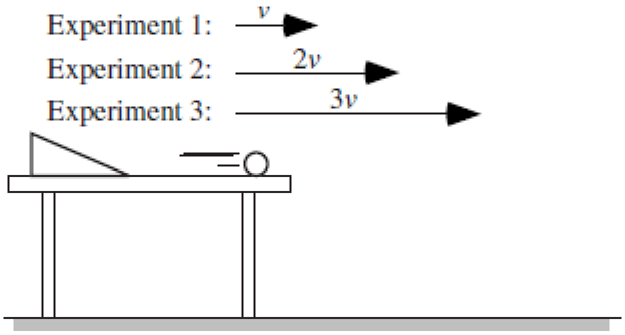


Figure 2.4: In 3 separate experiements, a student launches a ball horizontally from a table. The ball leaves the table with a speed v in the first experiment, with a speed $2v$ in the second experiment, and with a speed $3v$ in the third experiment. Rank, from greatest to least, the *time* that it takes the ball to travel from the table to the floor in the three experiments. Explain how you determined your ranking. This question is taken from Frank, Kanim, and Gomez (2008).

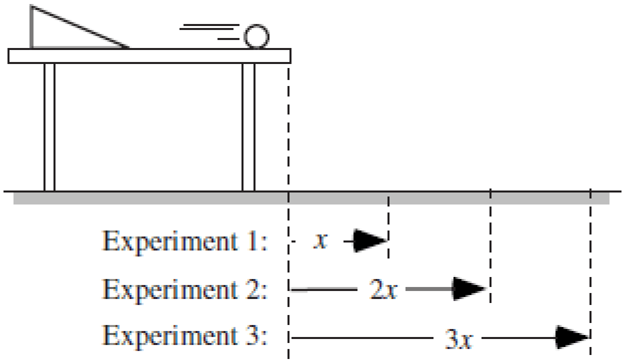


Figure 2.5: In 3 separate experiements, a student launches a ball horizontally from a table. The horizontal distance that the ball travels before hitting the ground is x in the first experiment, $2x$ in the second experiment, and $3x$ in the third experiment. Rank, from greatest to least, the *time* that it takes the ball to travel from the table to the floor in the three experiments. Explain how you determined your ranking. This question is taken from Frank, Kanim, and Gomez (2008).

How does a resources perspective address cognitive change? One idea is from diSessa and Sherin's (1998) work on what they call coordination classes. They note that most definitions of concepts and conceptual change are vague at best and absent at worst. In response, they define a specific type of concept, which they call a coordination class. A coordination class is a linked set of cognitive resources and the information as to when each should be activated, and its primary function is to determine how information from the external world should be interpreted and understood (diSessa and Sherin 1998). There are two types of change that this concept may experience. Conceptual change can occur with changes to the coordination class's readout strategies – which determine whether or not a given piece of information is important (diSessa and Sherin 1998). Conceptual change also occurs when the coordination class's cognitive resources and their relations to one another change (diSessa and Sherin 1998). In other words, conceptual change happens when any of the following are altered: resources, their relationships, and the contexts in which they are cued.

Do these resources-based studies completely invalidate the model of conceptual change posed by Posner *et al.* (1982)? Not necessarily. Strike and Posner (1992) acknowledge that what may, at first glance, appear to be a misconception is often not a robust mental structure present in the student's mind; it may very well result from the in-the-moment assembly of various pieces of knowledge. Nevertheless, they argue, the model proposed by Posner *et al.* (1982) is still valuable because it highlights the sort of evidence and experience people often require in order to experience conceptual change (Stike and Posner 1992). Furthermore, the model can help focus an instructor's attention on specific causes that may explain why a student struggles to experience conceptual change (Strike and Posner 1992). Thus, Posner *et al.*'s model of accommodation need not be discarded.

As a final perspective on conceptual change, consider the work of Vosniadou (1994). She proposes that people possess naïve framework theories which constrain how we acquire knowledge. Like diSessa's (1993) "sense of mechanism," these framework theories develop from our everyday observations of the world around us and we are often not consciously aware of their existence; unlike diSessa, Vosniadou argues that students *do* possess well defined models of phenomena, even if those models are first constructed on the spot in response to a particular context. Vosniadou further argues that framework theories contain many presuppositions about how the world should operate. Effective conceptual change requires that we target our instruction at these presuppositions and not on the mental models that a person generates on the spot when incoming information interacts with her framework theories (Vosniadou 1994). This requires making students conscious of their presuppositions and convincing them that they are not unassaible facts (Vosniadou 1994).

What lessons should one derive from these various models of cognition and conceptual change? First, conceptual change requires the (inter)active engagement of the learner. This appears to be true regardless of whether one adopts a misconceptions or a resources model. If conceptual change is to occur, then the student must accommodate a new concept (Posner *et al.* 2005), reorganize her resources and their cuing priorities and contexts (diSessa 1993; diSessa and Sherin 1998), and/or recognize the potential fallibility of her presuppositions (Vosniadou 1994). In other words, passively receiving information from an instructor is insufficient to effect conceptual change. Second, conceptual change involves a complex dance between what a student already knows and believes and what she is trying to learn and understand. In all the models considered above, conceptual change required the restructuring or even outright rejection of pre-existing mental patterns. This underscores why instruction must address students' prior knowledge in order to avoid emphemeral and superficial learning. Finally, instructors must know not only *what* their students are thinking but also *why* they are thinking it. How information is organized and accessed inside the mind matters, as the debate between the misconceptions and resources perspective emphasizes. Understanding a student's reasoning is important for helping them develop a dissatisfaction with a current concept (Posner *et al.* 2005), learn when to activate a certain resource (diSessa 1993), or verbalize their presuppositions (Vosniadou 1994). In the following section, I review some instructional strategies that reflect these ideas from the conceptual change literature.

2.2.3 Instructional Strategies

Given that students posses prior knowledge that influences how and what they learn and that may be highly resistant to traditional instruction, how can instructors address this knowledge in ways that promote more expert-like understandings? Two popular strategies for promoting conceptual change are *cognitive conflict* and *bridging*.

Cognitive conflict strategies derive heavily from the conceptual change model of Posner *et al.* (1982). They focus on creating dissatisfaction with a current idea and promoting the plausibility of a new idea. Posner *et al.* (1982) explicitly suggest such an approach when they argue that the teacher should confront students "with the problem arising from their attempts to assimilate new conceptions" and design "lectures, demonstrations, problems, and labs which can be used to create cognitive conflicts in students" (pp. 225-6). In the PER and AER literature, one of the most common cognitive conflict strategies is called *elicit-confront-resolve* (Heron 2004b; McDermott 1991). Each of these three verbs refers to a different step in the process. During the *elicit* stage, students are given a situation which evokes many common reasoning or conceptual errors. The students are then *confronted* with a new observation or line of reasoning that creates a conflict with their previous answer. They are then prepared to receive a new explanation that helps them *resolve* the conflict. A similar cognitive conflict strategy is called *observe-recognize-apply*

(McDermott 1991). In the *observe* stage, students make one or more observations or conduct one or more experiments. They are then asked to use their observations to answer one or more questions, during which they are lead to *recognize* the perhaps surprising implications of their observations. The students are then required to *apply* their new understandings to many different situations in order to prevent them reverting to their earlier, naïve ideas. Cognitive conflict strategies, such as the methods describe here, are, in many cases, associated with improved learning gains over traditional instruction (Heron 2004b; McDermott 1991; Prather *et al.* 2005).

While cognitive conflict has its advantages, it may not always be the best approach. If students feel as if they are constantly being "set up to fail," then the damage done to their attitudes and motivations may more than offset any potential learning gains (Redish 2003). Furthermore, as Strike and Posner (1992) note, if a student lacks a misconception on a topic, then attempting to "drown a students [*sic*] misconception into a 'sea of anomalies' may not be the best strategy" (p. 159) For these reasons, some instructors prefer an alternative approach known as bridging.

Instead of focusing on where students are wrong, bridging attempts to build upon situations in which their prior knowledge and intuitions are correct (Redish 2003). Clement, Brown, and Zietsman (1989) call such correct intuitions "anchoring conceptions." They recommend that whenever a student gives an incorrect answer, the instructor should propose a similar situation in which the student has an anchoring conception and is highly likely to give the correct answer (Brown and Clement 1989; Clement, Brown, and Zeitsman 1989). If the student does not accept the analogy between the two situations, then the instructor should propose one or more intermediate or bridging analogies that connect the two (Brown and Clement 1989; Clement, Brown, and Zeitsman 1989). For example, Brown and Clement (1989) describe how a student who believes a book experiences no force from the table on which it rests may change her thinking if she first considers a book resting on a spring and then thinks about a book resting on a pliable table. Unfortunately, not all bridging analogies are effective. Clement, Brown, and Zeitsman (1989) say that a book resting on one's hand does not help students imagine that a table can similarly exert an upward force on a book. They caution that some bridging analogies fail because they require students to alter some key physical aspect which, for the students, destroys the analogy between the current and target situations (Clement, Brown, and Zeitsman 1989). When effective, however, bridging analogies can be an effective means of promoting conceptual change (Clement, Brown, and Zeitsman 1989).

Elby (2001) adopts a similar approach to Clement, Brown, and Zeitsman (1989). His process, which he calls *refining intuitions*, is meant to improve students' attitudes and beliefs about science in addition to their knowledge of physics. One example of refining intuitions is his approach to teaching Newton's third law. He recognizes that many students think that, during a collision, a big truck exerts a larger force on a small car then the car exerts on the truck. He first guides his students through a series of exercises in which he helps students realize that the forces are equal. Next, he leads them through an activity in which helps them realize that their intuition that car has a "stronger reaction" than the truck is correct as long as "reaction" refers to acceleration, not force (Elby 2001). In this way, he shows students that "learning physics often involves *starting* with real-life example and common-sense intuitions, and building upon them to make careful decisions, to figure out equations, and so on" (Elby 2001, p. S62).

In principle, a series of carefully constructed items could lead a student through the processes of cognitive conflict and bridging without the student interacting with her peers. Yet many research-based curricular interventions and strategies emphasize student-student interactions. Why? Many instructors want to create students with the following characteristics:

> "They...continuously and actively probe their own understanding in the process of learning new concepts. They frequently formulate and pose questions to themselves, constantly testing their knowledge. They scrutinize implicit assumptions, examine systems in varied contexts, and are sensitive to areas of confusion in their understanding" (Meltzer and Manivannan 2002, p. 640).

Many students – especially novices in a discipline – do not yet possess these habits of mind. However, students may develop these habits when discussing their answers with their peers. Peer interactions force students to articulate their thinking. They also reduce the probability that a student will breeze through an activity with her naïve ideas intact. Finally, by explaining and defending their ideas to others (often in terms that are more accessible to other students than those used by their instructor), peer interactions help students realize when their ideas are insufficient or incorrect. A recent paper by Smith *et al.* (2009) provides evidence that peer discussions are responsible for helping students construct improved understandings of a topic.

2.3 Situated Models of Learning

2.3.1 Overview and Key Results

As Section 2.2 illustrates, cognitive models provide powerful explanations for how people learn. Yet many studies operating from the cognitive perspective make (often implicitly) what Greeno (2006) considers an invalid factoring assumption: Namely, a complex interacting system of people, tools, and environments may be understood by decomposing the system into its constituent elements, which may then be analyzed individually. But can we understand the behavior of an individual separate from her environment? Is the context in which action occurs merely the amalgamation of the separate activities of many individuals? Researchers utilizing the *situated perspective* answer "no" to both of these questions (Brown, Collins, and Duguid 1989; Finkelstein 2005; Greeno 1997; Greeno 2006; Vygotsky 1978). Where cognitive models focus on individuals, acquisition, and knowledge, situated models focus on contexts, participation, and knowing (Brown, Collins, and Duguid 1989; Finkelstein 2005; Greeno 1997; Greeno 2006; Sfard 1998).

The grammatical shift from nouns such as "knowledge" to adjectives such as "knowing" signifies an important distinction between situated and cognitive approaches to education (Sfard 1998). Cognitivists speak of "knowledge" as a thing that individuals acquire. Those adopting a situated perspective talk about "knowing," which is demonstrated by participation in culturally valued activities (Brown, Collins, and Duguid 1989; Greeno 1997; Greeno 2006; Sfard 1998). From a situated perspective, talking about "knowledge" that a person "possesses" is meaningless unless one also talks about the situations (hence, the *situated* perspective) in which such knowledge is evaluated (Greeno 2006). Furthermore, using an adjective such as "knowing" implies that learning is an active, ongoing process, deeply rooted in the contexts in which it occurs (Sfard 1998).

For example, Finkelstein (2005) describes the following incident from a class on teaching and learning physics. The students in this class were all upper-level undergraduates or graduate students. One of their assignments was designing a lesson plan to teach capacitance. One student was frustrated by this assignment: She felt confident that she understood capacitance until she was forced to explain it to others. After this student described her frustrations, the instructor wryly asked "What about writing a lesson plan could have made you forget what you already know?"

From the situated perspective, the issue is not about this student forgetting some piece of knowledge. Instead, the student's frustrations may be traced to the fact that her understanding of capacitance was being evaluated in a situation different from those she encountered when she was a student in introductory physics (Finkelstein 2005). Different situations (such as teaching capacitance versus using it to solve introductory physics problems) call for different levels of understanding. What one knows may be limited to and is often dependent on particular contexts (Finkelstein 2005; Greeno 2006).

From this perspective, learning is a process of enculturation (Brown, Collins, and Duguid 1989). That is, learning is manifested in one's increasing ability to participate in culturally valued activities and practices (Brown, Collins, and Duguid 1989; Greeno 1997; Greeno 2006; Sfard 1998). For those used to thinking about learning in terms of the acquisition of knowledge, this view of learning may seem strange. But consider the following argument from Brown, Collins, and Duguid (1989): There are many cases in which students can follow an algorithm, repeat a definition, or recite memorized information – yet they are totally incapable of applying these algorithms, definitions and bits of information in any meaningful way. Their minds may have acquired something, but, since they cannot use this information in any meaningful way, in what sense can we say they actually *know* these algorithms, definitions, and bits of information?

This point is underscored by Benezet's (1935a, 1935b, 1936) study of the teaching of arithmetic. Benezet was the superintendent of the Manchester, New Hampshire, schools during the Great Depression. In 1929, he instigated a radical restructuring of the mathematics curriculum: Until sixth or seventh grade, students received no formal instruction in arithmetic. Instead, teachers focused on developing students' skills in reasoning, estimation, and oral composition. Elementary students were only exposed to math to the extent it arose naturally in classroom discussions or when it related to some practical skill needed for life outside of the classroom (e.g. telling time, using money, and making measurements). Benezet found that students taught with this transformed curriculum could easily catch up with and even outperform their peers after a year or two of formal arithmetic in middle school.

What prompted Benezet to restructure the mathematics curriculum? From his classroom observations, he saw that students could do arithmetic in the sense that they could follow algorithms for addition, subtraction, multiplication, and division, and produce numbers. Yet he also saw many students who, when presented with a problem, would manipulate numbers with little thought as to why they should use a particular procedure or what a reasonable answer might look like (Benezet 1936). He also observed many students struggling to articulate mathematical ideas, such as why, when two fractions have the same numerator, the one with the largest denominator is the smaller of the two (Benezet 1935a). Finally, he worried that many elementary school students struggle with arithmetic simply because there are few contexts in which an elementary student needs arithmetic:

> "If I had my way, I would omit arithmetic from the first six grades. I would allow the children to practice making change with imitation money, if you will, but outside of making change, where does an eleven-year-old child ever have to use arithmetic?...What possible need has a ten-year-old child for a knowledge of long division?" (Benezet 1935a, p. 241).

Benezet's concerns make sense from the situated perspective: Prior to the restructuring, his students could not participate in all but the most restricted classrooom-based activities (Greeno 2006). Furthermore, his students' performances improved when they learned arithmetic in contexts that were commonly practiced by the broader community (Greeno 1997; Sfard 1998).

Does the situated perspective necessarily imply that education must always be conducted in complex, social environments, as some have claimed (Anderson, Reder, and Simon 1996)? According to Greeno (1997), not necessarily. Greeno argues that we should focus on instructional methods that enable participation discipline-valued activities. Acquiring facts, practicing routines, and individual exercises can all contribute to this goal. Thus, many traditional education practices are still valuable from the situated perspective, provided we view them as means toward increased participation and not ends unto themselves.

2.3.2 Implications for Instruction

What sort of pedagogical practices do situated models of learning emphasize? How does adopting a situated perspective refocus one's approach to teaching and assessment? In this section, I review some of the educational implications and recommendations that have emerged from the situated literature. By focusing on individuals' interactions with their environment, situated models of learning illuminate aspects of teaching, learning, and assessment that complement those highlighted by cognitive models (Sfard 1998).

Situated researchers stress that knowing is often distributed among many individuals, tools, and representations in a given environment. Greeno (2006), for example, talks about *distributed cognition* as the "problem solving, planning, and reasoning...accomplished by a group of people, working together with complex technological artifacts and with material representations they generate" (p. 84). Brown, Collins, and Duguid (1989) note that many people off-load some of the cognitive load associated with a task onto their environment. For example, a person who needed three-quarters of two-thirds of a cup of cottage cheese first measured out two-thirds of a cup, then placed it on a cutting board, and finally divided the cheese into fourths (Brown, Collins, and Duguid 1989). The Soviet psychologist Vygotsky, whose works have been highly influential in the situated community (Anderson, Reder, and Simon 1998; Shepard 2000), postulates that each person possesses a *zone of proximal development*; this marks the functions a person will soon be able to do by herself, but which she can now do with guidance from and collaboration with others (Vygotsky 1978). Ideas such as distributed cognition, task off-loading, and the zone of proximal development help explain why students learn more in classrooms that promote interactions with multiple tools, representations, and individuals. For these reasons, Finkelstein (2005) and Greeno (2006) note that situated models are often well-suited for analyzing the activities of collaborative classrooms.

Because learning, from a situated perspective, involves the participation in culturally-valued activities, situated researchers emphasize that students must learn both the tools and the culture of a discipline (Brown, Collins, and Duguid 1989; Sfard 1998; Shepard 2000). Instructors must provide students with practice at using the tools of a discipline in the same way as a practitioner of that discipline would. This means that assessments of student performance must be what Wiggins (1998) calls *authentic* (see also Shepard 2000). That is, they must "resemble the ways students will be expected to use their knowledge and skills in the real world" (Wiggins 1998, p. 4). Brown, Collins, and Duguid (1989) likewise define authentic activities as "the ordinary practices of the culture" (p. 34).

What happens when instructors do not use authentic assessments? McClymer and Knoles (1992) eloquently argue that inauthentic testing leads to ersatz learning. When instructors use inauthentic assessments, they establish a classroom culture in which students can succeed without knowing much about the culture of the discipline they are allegedly studying: "Students succeed because, despite the fact that the techniques and strageties they employ have only incidental connections to the disciplines involved, they are exceedingly well adapted to the sorts of tasks we ask them to perform in our courses" (McClymer and Knoles 1992, p. 34). Thus, students' learning is artificial. McClymer and Knoles highlight several techniques students use to succeed on inauthentic tasks, such as borrowing someone else's analysis and packing in data, jargon, and assertions. Students who can succeed by utilizing these techniques fail to become enculturated in a discipline and fail to achieve all but the most superficial forms of learning (McClymer and Knoles 1992).

2.4 Summary: Models of Learning, Interactive Engagement, and the Lecture-Tutorial Approach

This chapter began with an overview of the success of many interactive engagement strategies. Especially relevant for this dissertation is the success of the original *Lecture-Tutorials for Introductory Astronomy* (Prather *et al.* 2005; Prather *et al.* 2008). After reviewing two models of learning, the cognitive and the situated perspectives, we can now understand why interactive engagement generally and the *Lecture-Tutorials* specifically lead to higher learning gains than traditional forms of instruction.

Recall that interactive engagement strategies emphasize the construction of conceptual understanding through student-student and student-instructor discussion and feedback (Hake 1998a). From the cognitive perspective, the success of many interactive engagement strategies may be understood by the fact that they address, confront, and/or build upon students' prior knowledge, intuitions, and beliefs. Many interactive engagement strategies also help students organize their knowledge in more expert-like and easily retrievable ways. Finally, cognitive theories of conceptual change suggest where and how to target instruction. Many of the most popular interactive engagement resources to emerge from AER and PER have been designed and validated with these cognitive principles in mind (e.g. Heron 2004b; Hudgins *et al.* 2006; McDermott and Shaffer 1992b; Prather *et al.* 2005; Scherr 2002).

In particular, the *Lecture-Tutorials for Introductory Astronomy* were designed from a cognitive perspective (Prather *et al.* 2005). The topics of the *Lecture-Tutorials* were selected based on research into common student difficulties. Wherever possible, students' natural language was used instead of astronomy jargon. Questions were explicitly written to elicit and confront students' misconceptions. These design decisions explain, at least in part, the *Lecture-Tutorial*'s success at promoting increased student learning.

Situated models of learning also help explain the success of interactive engagement. First, the use of multiple representations, tools, and collaborations is concordant with situated models' focus on distributed cognition, task off-loading, and the zone of proximal development. Second, interactive engagement can foster student participation and enculturation in the discursive practices of a discipline (Swan 2008). Such enculturation is often an essential element of education.

Although the *Lecture-Tutorials* have not been extensively examined from a situated perspective (although Prather *et al.* 2005 do discuss implementation practices), the similar *Tutorials for Introductory Physics* (McDermott, Shaffer, and the Physics Education Group 2002) were given such a treatment by Finkelstein and Pollock (2005). They examined the various levels of contexts that support the successful implementation of these tutorials. Specifically, they note that the following elements of the course and institutional environments that are important for successful tutorial implementation:

(1) Students must make their ideas explicit to themselves and to one another, and instructors must elicit these ideas prior to the tutorial and assess students mastery of the tutorial's content post-tutorial;

(2) The tutorials must be well integrated into a set of activities and norms of the course that promote a consistent message about what learning physics involves;

(3) The department and/or college must provide adequate resources in terms of space, staff, and equipment to support the tutorials, which may be justified by the tutorials' role in improving students' educational experiences .

Finkelstein and Pollock's (2005) analysis demonstrates the various ways instructors, students, classroom environments and norms, and tutorials can interact to create a successful learning experience.

In this project, we follow the well-blazed path of the original *Lecture-Tutorials* and adopt a primarily cognitive perspective. Our focus is on effecting conceptual change in students in the realm of cosmology. This is why the cosmology lecture-tutorials employ cognitive change strategies such as cognitive conflict and bridging, as I discuss later. However, I will also use the situated perspective when it helps illuminate a particular aspect of lecture-tutorial success, lecture-tutorial implementation, and/or student learning.

Chapter 3

Selection of Topics

Modern cosmology is a broad field. A cursory glance at both introductory and advanced textbooks reveals a diverse - but related - set of topics. A list of these topics often includes the expansion of the universe, the Friedmann-Robertson-Walker (FRW) metric, the Big Bang, thermal relics, the Cosmic Microwave Background (CMB), Big Bang Nucleosynthesis, dark matter, dark energy, spatial curvature, inflation, Olber's paradox, and the formation of galaxies and large-scale structures (Cheng 2010; Dodelson 2003; Duncan and Tyler 2009; Kolb and Turner 1990; Peebles 1993). Because this list is so extensive, we deliberately focused our efforts on a subset of these topics.

This project examines common student difficulties with the expansion and evolution of the universe, the Big Bang, and the evidence for dark matter. These topics share two or more of the following characteristics:

(1) They are frequently taught in Astro 101 (see Slater *et al.* 2001 as well as introductory-level textbooks such as Duncan and Tyler 2009 and Bennett *et al.* 2008).

(2) They are conceptually complex, yet accessible to Astro 101 students (for example, we believe that Astro 101 students should be able to discuss the expansion of the universe and the Big Bang at the level of Lineweaver and Davis 2005).

(3) Previous studies indicate potential widespread difficulties and confusion (e.g. students' incorrect ideas about the Big Bang in Prather, Slater, and Offerdahl 2002).

The rest of this chapter specifies these topics in more detail, reviews pertinent prior research, and defends our choice of topics.

3.1 The Expansion of the Universe

Observations indicate that the universe is homogeneous and isotropic in space, but not in time. For example, the density of quasars in the universe (as measured in comoving volumes) increases as we look back in time (Sparke and Gallagher 2000). Additionally, observations are consistent with the theory that the CMB was once hotter in the past: CI absorption lines in a damped Lyman α system at a redshift of $z = 4.224$ are consistent with being excited by the CMB at a temperature of 14.2 K (Ledoux, Petitjean, and Srianand 2006), which is several degrees hotter than its currently measured temperature of 2.728 ± 0.004 K (Fixen *et al.* 1996). A universe that is homogeneous and isotropic in space but not in time is described by the FRW metric (Cheng 2010; Dodelson 2003; Kolb and Turner 1990; Peebles 1993). In spherical coordinates, the FRW metric is

$$ds^2 = -c^2 dt^2 + a^2(t) \left[dr^2 + f(r)(d\theta^2 + \sin^2\theta d\phi^2) \right] \ , \tag{3.1}$$

where

$$f(r) = \left\{ \begin{array}{c} \sin^2 r \\ r^2 \\ \sinh^2 r \end{array} \right\} \text{for a universe that is} \left\{ \begin{array}{c} \text{closed} \\ \text{flat} \\ \text{open} \end{array} \right\} . \tag{3.2}$$

Since current observations are consistent with a flat universe (e.g. Spergel *et al.* 2007), I will use $f(r) = r^2$ for the rest of this dissertation.

In order to satisfy Einstein's equation, the scale factor $a(t)$ must evolve according to the Friedmann equations:

$$\left(\frac{\ddot{a}}{a}\right) = -\frac{4\pi G}{3}(\rho + 3P) + \frac{\Lambda c^2}{3} \tag{3.3}$$

$$\left(\frac{\dot{a}}{a}\right)^2 = \frac{8\pi G\rho}{3} + \frac{\Lambda c^2}{3} \ . \tag{3.4}$$

In Equations (3.3) and (3.4), overdots represent time derivatives, G is the gravitational constant, Λ is the cosmological constant, and ρ and P represent the density and pressure, respectively, of the cosmological fluid. The density and pressure are also related to one another via the conservation equation (which one can derive from Equations (3.3) and (3.4)):

$$\dot{\rho} + 3\left(\rho + P\right)\frac{\dot{a}}{a} = 0 \; . \tag{3.5}$$

The equations show how the scale factor's evolution depend on the densities and pressures of the various components of the universe (radiation, matter, and dark energy).

At a constant coordinate time t, two points in the universe are separated by a radial distance d given by

$$d = \int ds = \int a(t)dr = a(t)r \; . \tag{3.6}$$

This distance d is often called the proper distance (e.g. Hogg 2000; Davis and Lineweaver 2004). Note that the coordinate separation r can remain constant while the proper distance grows if the scale factor increases with time t. This is why cosmologists say that, on average, galaxies remain fixed while space itself expands.

How do we know that the scale factor increases with time? Galaxies' redshifts provide the fundamental evidence. Mimicking the derivation in Sparke and Gallagher (2000), imagine a photon leaves a distant galaxy at a time t_e and arrives at Earth at a time t_r. This photon travels a coordinate distance r given by

$$r = \int_{t_e}^{t_r} \frac{cdt}{a(t)} \; . \tag{3.7}$$

If a second photon leaves the galaxy at a later time $t_e + \delta t_e$ and arrives at Earth at $t_r + \delta t_r$, then it travels a coordinate distance r given by

$$r = \int_{t_e+\delta t_e}^{t_r+\delta t_r} \frac{cdt}{a(t)} \; . \tag{3.8}$$

The coordinate distance traveled must be the same for both photons (assuming Earth and the galaxy did not change their coordinate locations), hence

$$\int_{t_e}^{t_r} \frac{cdt}{a(t)} = \int_{t_e+\delta t_e}^{t_r+\delta t_r} \frac{cdt}{a(t)} \; . \tag{3.9}$$

Equation (3.9) is true only if

$$\int_{t_e}^{t_e+\delta t_e} \frac{cdt}{a(t)} = \int_{t_r}^{t_r+\delta t_r} \frac{cdt}{a(t)} \tag{3.10}$$

is also true. If δt_e and δt_r are small such that the scale factor is approximately constant in each integral in Equation (3.10), then the integrals evaluate to

$$\frac{\delta t_e}{a(t_e)} = \frac{\delta t_r}{a(t_r)} \tag{3.11}$$

or

$$\frac{\delta t_r}{\delta t_e} = \frac{a(t_r)}{a(t_e)} \ . \tag{3.12}$$

If, instead of two photons, we imagine we are dealing with two crests of an electromagnetic wave, then Equation (3.12) becomes

$$\frac{a(t_r)}{a(t_e)} = \frac{\nu_e}{\nu_r} = \frac{\lambda_r}{\lambda_e} = 1 + z \ , \tag{3.13}$$

where the νs reprsent frequencies, the λs represent wavelengths, and z is the redshift. The final equality in Equation (3.13) is simply the definition of redshift. Equation (3.13) shows that the redshift is positive when the scale factor grows over time and is negative (i.e. is a blueshift) when the scale factor shrinks over time. Since the early twentieth century, astronomers have observered that most galaxies have redshifts proportional to their distances (Hubble 1929), thus indicating that the scale factor has grown.

The proportionality between redshifts and distances is a manifestation of Hubble's law. Hubble's law is traditionally written as

$$v_{rec} = H(t)d \ , \tag{3.14}$$

where v_{rec} represents recession velocity and $H(t)$ is Hubble's parameter. Note that the recession velocity is related to redshift through the general relativistic relation given in Davis and Lineweaver (2004). Hubble's law emerges naturally from differentiating Equation (3.6) with respect to time:

$$\frac{dd}{dt} = \frac{da}{dt}r + a\frac{dr}{dt} \ . \tag{3.15}$$

The second term on the right hand side of Equation (3.15) represents the peculiar velocity of a galaxy (Davis and Lineweaver 2004). The first term is the recession velocity (Davis and Lineweaver 2004). Since $H(t) = \frac{da/dt}{a}$, this term is equivalent to Hubble's law.

Plots of recession velocities versus distances are known as Hubble plots. The slope of a Hubble plot is determined by Hubble's parameter. If $H(t)$ is constant in time (as is the case for a universe expanding at a constant rate), then the Hubble plot is a straight line with a positive slope. If $H(t)$ changes with time, then the slope of the Hubble plot will change with distance (since farther distances correspond to looking further back in time). Astronomers have known that $H(t)$ is changing with time since the discovery of dark energy and the accelerating expansion of the universe (Perlmutter *et al.* 1999; Riess *et al.* 1998).

The idea that the universe is expanding and the observations supporting this idea are fundamental to modern cosmology. Yet previous studies indicate that they are neither well-known nor well-understood. Lightman, Miller, and Leadbeater (1987) and Lightman and Miller (1989) report that most people are unaware that the universe is expanding and, in absence of any evidence to the contrary, presume it is static. Lightman, Miller, and Leadbeater (1989) further report that 43% of the high school students they surveyed expressed fear or other negative psychological associations with the prospect of an expanding universe. Furthermore, Prather, Slater, and Offerdahl (2003) and Lineweaver and Davis (2005) describe how people think of the Big Bang as an explosion of (pre-existing) matter into empty space, as opposed to the beginning of the expansion of space. Finally, although they are not specifically related to cosmology, the studies of Trowbridge and McDermott (1980; 1981) on student difficulties with velocity and acceleration and the study of McDermott, Rosenquist, and van Zee (1987) of student difficulties with graph interpretation suggest that students might have trouble understanding Hubble plots. Taken together, these studies suggest fundamental deficiencies in people's knowledge and reasoning about the expansion of the universe.

3.2 The Big Bang

Like the closely-related topic of expansion, evidence suggests that the Big Bang is widely misunderstood. As mentioned in Section 3.1 above, many think of the Big Bang as an explosion of (pre-existing) matter into empty space (Prather, Slater, and Offerdahl 2003; Lineweaver and

Davis 2005). Prather, Slater, and Offerdahl (2003) hypothesize that students' belief in pre-existing matter may be a manifestation of a p-prim "you can't make something from nothing." Prather, Slater, and Offerdahl (2003) as well as Simonelli and Pilachowski (2004) also note that a sizable minority of students talk about the Earth, Solar System, and formation of the Solar System when discussing the Big Bang.

A more accurate formulation of the Big Bang Theory is that it describes the evolution of the universe from an initially hot and dense state. Note that this does not mean that the entire universe was once small, as Lineweaver and Davis (2005) are careful to stress:

> "The ubiquity of the big bang holds no matter how big the universe is or even whether it is finite or infinite in size. Cosmologists sometimes state that the universe used to be the size of a grapefruit, but what they really mean is that the part of the universe we can now see – our observable universe – used to be the size of a grapefruit....[W]e can conceive of the early universe as a pile of overlapping grapefruits [representing the observable universes of different observers] that stretches infinitely in all directions. Correspondingly, the idea that the big bang was 'small' is misleading. The totality of space could be infinite. Shrink an infinite space by an arbitrary amount, and it is still infinite" (p. 40).

Indeed, if the universe is spatially flat, then its volume must be infinite (assuming the FRW metric holds for all space). When one integrates the FRW metric for a flat universe over all space one obtains an infinite answer. This result, as well as the idea of overlapping observable universes, also refutes the idea that the universe has a center and an edge (assuming that the universe really is homogeneous and isotropic in space).

Even though the universe was not "small" early in its history, densities were still higher than they are today. This can be seen from Equation (3.5). For matter, which is pressureless ($P_m = 0$), Equation (3.5) becomes

$$\frac{d}{dt}\left(\rho_m a^3\right) = 0 \ , \tag{3.16}$$

which implies the density of matter ρ_m scales as a^{-3}. For radiation ($\rho_r = 3P_r$), Equation (3.5) becomes

$$\frac{d}{dt}\left(\rho_r a^4\right) = 0 \ , \tag{3.17}$$

and so the density of radiation ρ_r scales as a^{-4}. For dark energy ($\rho_\Lambda = P_\Lambda$) the continuity equation is simply

$$\frac{d}{dt}(\rho_\Lambda) = 0 \ , \tag{3.18}$$

so ρ_Λ is constant over the history of the universe. These scaling relations show that at earlier times, when the scale factor was smaller, the densities of matter and radiation were larger than they are today.

Likewise, temperatures were hotter in the early universe. The density ρ and temperature T of a relativistic fluid is given by

$$\rho = \frac{g\pi^2}{30}T^4 \ , \tag{3.19}$$

where g is the effective number of relativistic species (Dodelson 2003). The temperature of the universe is dominated by CMB photons, whose density scales as a^{-4}, as shown above. Thus, the temperature of the universe must scale as a^{-1}.

This scaling of the temperature with the scale factor indicates that, at early times, the universe's temperature was high enough that the average particle energy was much greater than the energy needed to create a single baryon, such as a proton. During these times, photon collisions could easily produce particle-antiparticle pairs, which in turn would collide to create photons. These reactions were initially in thermal equilibrium. However, as the universe expanded and the temperature dropped, the thermal energy fell and eventually the reaction rates became slower than the expansion rate of the universe. At this point, photons were no longer converted into matter and antimatter. The remaining antimatter and much of the remaining matter annihilated. The fact that a residual amount of matter was not annihilated in matter-antimatter interactions implies there was an initial asymmetry between the amount of matter and the amount of antimatter created in the early universe. The reason for this asymmetry is the subject of studies on baryogenesis and involves particle physics beyond the scope of this overview (see Kolb and Turner 1990 for more information).

How is this sequence of events typically presented to Astro 101 students and other non-experts? Most non-technical treatments simply say that energy was converted into matter. For example, Duncan and Tyler (2009) write

> “The very early universe is difficult to imagine because the conditions were so unlike what we experience today. Enough energy was packed into each bit of space that many different kinds of particles could freely pop in and out of existence and transform from one particle into another” (pp. 327-8).

Hawking, in his popular book *A Brief History of Time*, makes a similar, if somewhat longer, argument:

> “There are something like ten million million million million million million million million million million million million million million (1 with eighty zeroes after it) particles in the region of the universe we can observe. Where did they all come from? The answer is that, in quantum theory, particles can be created out of energy in the form of particle/antiparticle pairs. But that just raises the question of where the energy came from. The answer is that the total energy of the universe is exactly zero. The matter in the universe is made out of positive energy. However, all matter is attracting itself by gravity. Two pieces of matter that are close to each other have less energy than the same two pieces a long way apart, because you have to expend energy to separate them against the gravitational force that is pulling them together. Thus, in a sense, the gravitational field has negative energy. In the case of a universe that is approximately uniform in space, one can show that this negative gravitational energy exactly cancels the positive energy represented by the matter. So the total energy of the universe is zero” (pp. 128-9).

Such treatments also frequently invoke Einstein’s famous equation $E = mc^2$. For example, the textbook by Bennett *et al.* (2008) makes the following statement:

> “The universe was so hot during the first few seconds that photons could transform themselves into matter, and vice versa, in accordance with Einstein’s formula $E = mc^2$. Reactions that create and destroy matter are now relatively rare in the universe at large, but physicists can reproduce many such reactions in their laboratories” (p. 707).

As described in subsequent chapters, our lecture-tutorials adopt a similar treatment to those presented above of the Big Bang and the evolution of the universe.

3.3 Expansion's Effects on Lookback Times and Distances

The expansion of the universe complicates our understanding of distances and lookback times. Astronomers often say that when one looks at an object X light-years away, one sees how that object appeared X years ago. While this simple relationship works well for objects in the nearby universe, it fails for objects at cosmological distances for the simple fact that objects are receding from us even as their light travels toward Earth. As Davis and Lineweaver (2004) and Lineweaver and Davis (2005) explain in more and less technical detail, respectively, the farthest object we can see is currently about 46 billion light-years away from us (and receding away many times faster than the speed of light), even though we cannot see what the universe looked like more than 14 billion years ago (since the universe is only about 13.7 billion years old; Spergel *et al.* 2007). Conceptual treatments of this issue can be found in introductory-level textbooks (e.g. Bennett *et al.* 2008) as well as articles written for the general public (Lineweaver and Davis 2005).

Consider the following example: A photon leaves a distant galaxy when the universe is 5 billion years old. The photon arrives at Earth 8 billion years later. What was the distance to the galaxy when the photon left? What was the distance to the galaxy when the photon arrived?

As Equation (3.7) shows, to calculate r we must first know how the scale factor changes with time. The evolution of the scale factor has, for much of the history of the universe, been driven by the density of matter in the universe (Sparke and Gallagher 2000). For a matter dominated universe, $a(t) \sim t^{2/3}$ (Sparke and Gallagher 2000). Thus,

$$\frac{a(t)}{a_0} = \left(\frac{t}{t_0}\right)^{2/3} , \tag{3.20}$$

where the subscript 0 indicates values measured "today" (which corresponds to an age of 13 billion years in this situation). Integrating from the time of emission t_e to the time of reception t_r yields

$$r = \frac{3ct_0^{2/3}}{a_0}\left[t_r^{1/3} - t_e^{1/3}\right] . \tag{3.21}$$

Setting $a_0 = 1$ gives a distance of 11 billion light-years. This is the current distance between the galaxy and Earth.

To find the distance between the galaxy and Earth when the light was emitted, we must know what the scale factor was when the universe was 5 billion years old. Since we set $a(t = 13 \text{ Gyr}) = 1$, we can use Equation (3.20) to find $a(t = 5 \text{ Gyr}) = 0.53$. This means the distance, according to Equation (3.6), was 5.6 billion light-years when the universe was only 5 billion years old.

In this situation, the photon traveled for 8 billion years, even though the galaxy was 5.6 billion light-years from Earth when the photon left and 11 billion light-years from Earth when the photon arrived. This situation illustrates why astronomers prefer to talk about lookback times and redshifts as opposed to measures of distance.

3.4 Dark Matter

Observations indicate that most of the matter in the universe is composed of non-baryonic particles. The evidence for this so-called dark matter rests on a concordance of observations of its gravitational effects. The effects of dark matter are observed in a variety of settings, including the CMB's power spectrum (Spergel *et al.* 2007), the rotation curves of spiral galaxies (Rubin, Ford, and Thonnard 1980), galaxies' motions in clusters (Zwicky 1937), and cluster X-ray gas (Lewis, Boute, and Stocke 2003). Although the evidence for dark matter is often addressed at several points in an Astro 101 course, most students first encounter it when they study the rotation curves of spiral galaxies (see, for example, the introductory textbooks of Duncan and Tyler 2009 and Bennett *et al.* 2008).

A galaxy's rotation curve is a plot of orbital velocities as a function of distance from the its center. Such plots are made from spectroscopic observations of Hα emission lines, CO rotational transition lines, and maser lines, to name a few (for more lines of interest and for an overview of how one uses these lines to produce rotation curves, see Sofue and Rubin 2001). Rotation curves provide important information on the dynamics, evolution, and distribution of mass in spiral galaxies (Sofue and Rubin 2001).

In the 1970s, astronomers first noticed that spiral galaxies have rotation curves that remain flat at large distances from their galactic centers (Rubin, Ford, and Thonnard 1980). This was a surprise, because the density of luminous matter appears to fall off as one moves away from the center of a galaxy. Astronomers expected that, at large radii, rotation curves should show a Keplerian fall-off (that is, the velocities should drop as $r^{-1/2}$) since most of the luminous matter should be enclosed by circular orbits at these radii. Since the velocity v of an object orbiting a galaxy scales as $\sqrt{M/r}$ (where M represents the mass interior to that object's orbit), the only way a galaxy's rotation curve can remain flat is if $M \sim r$ – or, in other words, if there is more mass in the galaxy than we observe emitting light. (There is another possibility, on which I will not dwell: We may need to reformulate our theories of gravity. This is currently a minority opinion within the astronomical community. See Sanders and McGaugh 2002 for more details.)

While this chain of reasoning is normally presented to Astro 101 students to convince them that dark matter exists, there has been no research (to our knowledge) on what students actually take away from such a presentation. One may suspect that students struggle with this information since rotation curves were once counterintuitive even to professional astronomers. Furthermore, this approach assumes that Astro 101 students can interpret the graphical information encoded in rotation curves – which, as McDermott, Rosenquist, and van Zee (1987) demonstrate, may not be a safe assumption. Whether or not students understand why rotation curves are evidence for dark matter is an open question for this study.

3.5 Summary: Justifying the Cosmology Lecture-Tutorial Topics

Expansion, the Big Bang, and dark matter are certainly not the only cosmology topics worthy of study by astronomy education researchers. However, they capture many of the key aspects of cosmology, such as the evolution, fate, and composition of the universe. They also present multiple opportunities for Astro 101 students to interpret graphs (such as Hubble plots and rotation curves), explore the strengths and limitations of scientific analogies (such as the balloon analogy for expansion), and make inferences from data (e.g. by using a Hubble plot to infer whether or not the expansion rate changes or by using a rotation curve to infer the presence of dark matter in spiral galaxies). These topics thus provide multiple opportunities for students to practice the kinds of reasoning skills many Astro 101 instructors explicitly or implicitly strive to teach.

These topics also present numerous opportunities for students to make conceptual or reasoning errors. Some of these difficulties have been previously documented in the research literature (Lightman, Miller, and Leadbeater 1987; Lightman and Miller 1989; Comins 2001; Prather, Slater, and Offerdahl 2003; Davis and Lineweaver 2004; Lineweaver and Davis 2005; Simonelli and Pilachowski 2004). Some are hinted at by studies on related and foundational topics (e.g. Trowbridge and McDermott 1980; Trowbridge and McDermott 1981; McDermott, Rosenquist, and van Zee 1987). Others were uncovered for the first time by this study (see Chapters 6-8). The presence of these conceptual and reasoning difficulties suggests that curricular interventions based on research-supported models of learning (see Chapter 2), such as lecture-tutorials, may help many students overcome these stumbling blocks in learning cosmology.

Chapter 4

Design of Surveys

The process of studying students' difficulties with the expansion of the universe, the Big Bang, and dark matter, and designing lecture-tutorials to help them overcome their difficulties falls under Schoenfeld's (2009) definition of a *design experiment*:

> "Properly constructed, a design experiment consists of the creation of an instructional intervention on the basis of a local theory regarding the development of particular understandings. The intervention is then examined with regard to the accuracy of the underlying local theory *and* the power of the intervention, with an eye toward refining both" (p. 3, italics in original).

Schoenfeld (2008) calls attention to the design aspect of this process by noting that such experiments "would be improved if the investigators could rely on accessible and useful principles of design" (p. 2). Others have also highlighted the importance of the design process in developing curricular interventions (e.g. Schunn 2008; Swan 2008). Schunn (2008), for example, argues that the design process can have a significant influence over the success or failure of the intervention since, as time goes on, one's flexibility in possible designs decreases and the cost of redesign increases.

For this study, design issues are not limited to the development of the lecture-tutorials. We also had to develop our own surveys of conceptual cosmology knowledge. Currently, there exist surveys designed to measure students' knowledge of lunar phases (Lindell 2001), the greenhouse effect (Keller 2006), light and spectroscopy (Bardar *et al.* 2007), and star properties (Bailey 2007), as well as the more general Astronomy Diagnostics Test (Hufnagel 2001); yet no comperable survey of students' conceptual cosmology knowledge exists, no doubt due to the lack of prior research on this subject. In this chapter, I discuss the design of four conceptual surveys of cosmology topics. I will describe the design of the cosmology lecture-tutorials in Chapter 5.

There are many well-articulated principles for effective designs (Schunn 2008; Wilson 2005). For this project, we follow the four step process of survey design and interpretation recommended by Wilson (2005):

(1) Define the constructs to be measured and create construct maps for each construct;

(2) Design survey items;

(3) Score and categorize the full range of responses; and

(4) Apply psychometric measurement models to the data.

This chapter focuses on the first two steps. The last two are considered in more detail in Chapters 6-8.

4.1 Constructs and Construct Maps

In this dissertation, the term *construct* refers to "the concept or characteristic that a test is designed to measure" (AERA, APA, and NCME 1999, p. 5; see also Wilson 2005). There is widespread agreement that developing a robust and specific definition of a construct is always the first step one should take in survey development (Gorin 2006; Shepard 1993; Wilson 2005). By defining the construct of interest in advance, survey designers can select items that accurately probe multiple levels and/or multiple components of the construct (Gorin 2006; Shepard 1993; Wilson 2005).

Unfortunately, one can easily write a construct definition that is too vague to be useful. Gorin (2006) describes why such vague definitions place survey designers in a poor position:

> "Items on these tests are traditionally written so that each item can be tied to at least one of the standards to be measured. However, whether phrased as a verbal definition or as a list of standards-based skills, the generality of their language presents a significant limitation for test development and validation. In terms of

> item writing, how can an item writer efficiently develop tasks without an understanding of the various skills comprising a domain or a curriculum standard? In terms of validation, how can evidence be gathered to support inferences about cognition when no cognitive terms have been defined?" (p. 21).

The key to good survey design is to operationalize, to the greatest extent possible, what exactly the survey is supposed to measure.

One way to help operationalize what a survey measures is to create a *construct map* for each survey. According to Briggs *et al.* (2006),

> "Construct maps are used to represent unidimensional continua with distinct levels. Each level reflects a heirarchical stage through which students pass as they gain a qualitively richer understanding about a given construct" (p. 38).

The creation of a construct map thus depends on whether people can be placed in an ordered hierarchy by how much of the construct they "possess" (Wilson 2005). The construct map itself is a visual or tabular instantiation of a construct and how people may vary on that construct (Wilson 2005). For examples of construct maps, see Briggs *et al.* (2006), Wilson (2005), and Tables 4.1-4.4 below.

Why does this process of construct definition and construct map creation matter? Because the validity of a survey – that is, whether or not the survey actually measures what its designers believe it measures – depends on the process by which items are selected for inclusion or exclusion from the survey (Briggs *et al.* 2006). Part of the validity argument for a survey depends on establishing whether that the survey's items adequately cover the survey's construct (Gorin 2006; Shepard 1993). Thus, the process underlying the construction of the survey is a necessary (although not sufficient – see, for example, Shepard 1993) condition for establishing the validity of the survey.

As one may guess after reading Chapter 3, this study encompasses multiple constructs. Specifically, we defined four constructs of interest:

(1) interpreting Hubble plots;

(2) models of the expansion of the universe and the Big Bang;

(3) the evolving universe; and

(4) evidence for dark matter in spiral galaxies.

These constructs are hereafter referred to by their abbreviated names *Hubble plots*, *models*, *evolving universe*, and *dark matter*, respectively. They are defined and described in more detail below, along with their associated construct maps.

Table 4.1: The construct map for the *Hubble plots* construct.

Level	Description
3	The student **correctly** reasons about the age and the expansion rate of the universe using Hubble plots.
2	The student **sometimes correctly** reasons about Hubble plots, but **sometimes cues on the wrong features of the graph or incorrectly interprets a feature**.
1	The student **incorrectly** reasons about the age and the expansion rate of the universe using Hubble plots.

The *Hubble plots* construct looks at whether or not a student can use Hubble plots to reason about the age and expansion rate of the universe. Since this construct obviously relies on students' abilities to interpret graphs, we looked at McDermott, Rosenquist, and van Zee's (1987) work on student difficulties in reading graphs. McDermott, Rosenquist, and van Zee (1987) found that students often focus on an inappropriate feature of a graph (for example, they might look at the height of a graph when they need to focus on its slope or vice-versa). Students also struggle to interpret negative quantities, such as velocity and acceleration. Finally, they found that students tend to think that the graph of an object's motion should resemble the path of its motion (e.g. a ball rolling down a hill should have a velocity versus time graph and/or an acceleration versus time graph that is/are shaped like the hill; see also Elby 2000). We hypothesized that Astro 101 students encounter similar difficulties reading Hubble plots, especially because Hubble plots include information on kinematic quantities such as velocity (Trowbridge and McDermott 1980) and acceleration (Trowbridge and McDermott 1981).

Table 4.1 is the construct map for the *Hubble plots* construct. This construct map represents a hypothesis about the different levels of mastery into which students may fall on this construct.

Table 4.2: The construct map for the *models* construct.

Level	Description
4	The student **correctly** states that the universe is physically expanding in size over time. The student **correctly** states that only galaxies are moving apart from one another due to expansion. The student **correctly** claims that the universe has no center. The student **correctly** claims that the universe has no edge. The student **correctly** describes the Big Bang as the beginning of expansion.
3	The student **correctly** states that the universe is physically expanding in size over time. The student **incorrectly** states that all objects in the universe are moving apart from one another due to expansion. The student **correctly** claims that the universe has no center. The student **correctly** claims that the universe has no edge. The student **may** describe the Big Bang as an explosion or as the beginning of expansion.
2	The student **correctly** states that the universe is physically expanding in size over time. The student **incorrectly** states that all objects in the universe are moving apart from one another due to expansion. The student **incorrectly** claims that the universe has a center. The student **incorrectly** claims that the universe has an edge. The student **incorrectly** describes the Big Bang as an explosion but not as the beginning of something smaller than the universe (e.g. the Solar System, Galaxy, etc.).
1	The student **incorrectly** states that the universe is not physically expanding in size over time. The student **may or may** not claim that the universe has a center. The student **may or may** not claim that the universe has an edge. The student **incorrectly** describes the Big Bang as an explosion and/or as the beginning of something smaller than the universe (e.g. the Solar System, Galaxy, etc.).

At the lowest level (Level 1) are students that always incorrectly interpret Hubble plots. At the highest level (Level 3) are students that correctly use Hubble plots to qualitatively reason about the expansion rate and age of the universe. Students who sometimes correctly reason using Hubble plots fall in the middle region (Level 2). We hypothesized that students who fall in Level 1 or Level 2 make many of the graph interpretation errors outlined by McDermott, Rosenquist, and van Zee (1987). Additionally, we hypothesize that students in Levels 1 and 2 may also neglect that the farther away one looks in the universe, the further back in time one sees. This fact adds an astronomical twist to the difficulties discussed in McDermott, Rosenquist, and van Zee (1987). In Section 4.3, I describe how we designed items to help place students at their appropriate levels on this construct map.

The *models* construct focuses on students conceptualizations of the expansion of the universe and the Big Bang. Much of the previous research on conceptual difficulties with cosmology applies to this construct (Comins 2001; Lightman, Miller, and Leadbeater 1987; Lightman and Miller 1989; Prather, Slater, and Offerdahl 2003; Simonelli and Pilachowski 2004). These previous studies influenced our design of the construct map for this construct.

Table 4.3: The construct map for the *evolving universe* construct.

Level	Description
3	The student **correctly** relates the light travel time between two galaxies to their past, present, or future distances from one another. The student **correctly** describes how the temperature of the universe has changed over time. The student **correctly** describes how the density of matter in the universe has changed over time. The student **correctly** states that the matter in the universe has **not** always existed.
2	The student **correctly** relates the light travel time between two galaxies to their past, present, or future distances from one another. The student **correctly** describes how the temperature of the universe has changed over time. The student **correctly** describes how the density of matter in the universe has changed over time. The student **incorrectly** states that the matter in the universe has always existed.
1	The student **incorrectly** relates the light travel time between two galaxies to their past, present, or future distances from one another. The student **incorrectly** describes how the temperature of the universe has changed over time. The student **incorrectly** describes how the density of matter in the universe has changed over time. The student **incorrectly** states that the matter in the universe has always existed.

The *models* construct map is shown in Table 4.2. People who do not know that the universe is expanding or that the Big Bang is related to the expansion of the universe are at the lowest level

(Level 1). At Level 2 are people who conceive of expansion and the Big Bang as the motion of objects in the universe away from a center and into empty space. At the highest levels (Levels 3 and 4) are those who relate the expansion of the universe and the Big Bang to the expansion of space itself. People at Level 3 are almost identical to people at Level 4 with one exception: Those at Level 3 erroneously claim that all objects in the universe – planets, stars, galaxies, etc. – move away from one another due to the expansion of the universe. In general, only the distances between galaxies are affected by the expansion of the universe. Although this construct map is based in part on prior studies, I must emphasize that it, like all the construct maps presented here, originally represented a hypothesis about how students are arranged along this construct.

Table 4.4: The construct map for the *dark matter* construct.

Level	Description
4	The student **correctly** identifies the rotation curve of a spiral galaxy. The student **correctly** describes how the orbital speeds of stars at different radii relate to one another based on the rotation curve s/he chose. The student **correctly** describes how mass is distributed in the galaxy based on the rotation curve s/he chose.
3	The student **incorrectly** identifies the rotation curve of a spiral galaxy. The student **correctly** describes how the orbital speeds of stars at different radii relate to one another based on the rotation curve s/he chose. The student **correctly** describes how mass is distributed in the galaxy based on the rotation curve s/he chose.
2	The student **incorrectly** identifies the rotation curve of a spiral galaxy. The student **correctly** describes how the orbital speeds of stars at different radii relate to one another based on the rotation curve s/he chose. The student **incorrectly** describes how mass is distributed in the galaxy based on the rotation curve s/he chose.
1	The student **incorrectly** identifies the rotation curve of a spiral galaxy. The student **incorrectly** describes how the orbital speeds of stars at different radii relate to one another based on the rotation curve s/he chose. The student **incorrectly** describes how mass is distributed in the galaxy based on the rotation curve s/he chose.

The *evolving universe* construct looks at whether or not a student knows how properties of the universe have changed over time. Specifically, this construct focuses on students' knowledge of how expansion has affected the amount of matter in the universe, the density of matter in the universe, the temperature of the universe, and the relationship between lookback times, proper distances, and light travel time between widely-separated galaxies. The *evolving universe* construct map is shown in Table 4.3.

The *dark matter* construct probes whether a student can construct the causal chain of reasoning linking the flat rotation curves of spiral galaxies to the existence of dark matter. The construct map for this construct is displayed in Table 4.4.

The construct map in Table 4.4 has four levels. At the lowest level are people who do not demonstrate any correct link in the chain of reasoning described above. At Level 2 are people who, despite selecting an incorrect rotation curve, correctly relate the orbital speeds of stars at various radii using that rotation curve. If someone can also connect the orbital speeds of stars to the distribution of mass in the galaxy, then she will be at Level 3. Level 4 is reserved for students who select the right rotation curve and correctly relate it to the orbital speeds of stars and the distribution of mass in the galaxy. We hypothesize that some students will pick the wrong rotation curve but then correctly use that rotation curve to connect the orbital speeds of stars and and the distribution of mass in the galaxy. Since connecting these three ideas may be non-trivial for the average Astro 101 student, a student who can make such a connection should be placed high on the construct map. After all, she may understand the relevant physics – she just does not know what the true rotation curve of a spiral galaxy looks like, much like astronomers several decades ago.

Given these constructs and construct maps, how can one design items that allow students to be placed at the appropriate positions on the construct maps? I describe the principles underlying the design of our items in the following section.

4.2 Principles Guiding Design of Items

In some respects, the process of designing items involves a certain amount of creativity or inspiration of the survey designer (Wilson 2005). However, such “item brainstorming sessions” must be complemented by a detailed evaluation of whether or not the items adequately cover the construct of interest (Shepard 1993). Furthermore, there are potentially an infinite (or at least a very large) number of items one could include in the survey. Wilson (2005) describes the process of item design as selecting items from the pool of potential items:

> "One way to understand the *items design* is to see it as a description of the population of items, or 'item pool,' from which the specific items in the instrument are to be sampled. As such the instrument is the result of a series of decisions that the measurer has made regarding how to represent the construct or, equivalently, how to stratify the 'space' of items (sometimes called the *universe* of items) and then sample from those strata. Some of those decisions are principled ones relating the fundamental definition of the construct and the research background of the construct. Some are practical, relating to the constraints of adiminstration and usage. Some are rather arbitrary, being made to keep the item-generation task within reasonable limits" (pp. 44-5, italics in original).

He goes on to define two key components of items that help determine the item pool: 1) the construct component and 2) the descriptive component.

The construct component of a survey's items reflects the degree to which they help place respondents at various levels of the construct (Wilson 2005). When designing an item, one should always ask "If a student answers this item correctly or incorrectly, what does that tell me about the amount of the construct she 'possesses'?" The construct map associated with a construct should specify the attributes a survey's items need to elicit from respondents (Wilson 2005).

The descriptive component refers to all the other required properties of the items that do not relate to the construct (Wilson 2005). For example, should the survey include only free-response items, only multiple choice items, or a combination of both? How quickly can the average student provide a complete and correct response? Is the use of jargon essential or superfluous? One must answer these and related questions when reducing the item pool.

For the four surveys we wrote (one for each construct, as described in Sections 4.3-4.6 below), we were cognizant of several constraints to the descriptive component of our items. First, the majority of items have to be open-ended. Open-ended questions have the potential to reveal common reasoning difficulties among students as well as students' natural language; this is why many multiple-choice tests begin as free-response questions early in their development phase (e.g. Bailey 2007; Bardar et al. 2007). Given the lack of research on Astro 101 students' struggles with cosmology, we simply do not have enough information to create effective multiple choice items. However, asking too many open-ended questions quickly leads to student fatigue and a corresponding degradation in item responses. Additionally, instructors are generally only willing to give up about twenty minutes of class time for a survey. One solution is to provide students with response options (thereby limiting the amount of time they need to indicate their answers) while still asking them to explain why they chose their answer. This is the approach we took with a minority of items. Furthermore, we limited the number of items on each survey. In general, each survey contained one item per attribute on its associated construct map. While more items would undoubtedly increase the accuracy and precision with which we place a student at a given level on a construct map, they would also lengthen the test and reduce the efficacy of individual questions.

Our survey items are subject to other constraints as well. In order to use the surveys for pre-post testing, the questions have to be worded using language that is accessible to students regardless of whether or not they have received instruction. Yet the question must still somehow probe a student's understanding. We here adopt Heron's (2004a) definition of understanding: A student understands a topic "if, when faced with an unfamiliar problem, he or she reliably selects the appropriate concepts and principles, applies them correctly, and constructs a logically sound solution" (p. 342). We could not restrict our item pool to items that simply elicit factual recall since students may correctly answer such items without understanding the underlying concepts (Vosniadou 1994). This suggested creating questions modeled after what Vosniadou (1994) calls generative questions. A generative question confronts people "with phenomenon about which they do not have any direct experience and about which they have not yet received explicit instruction. Because generative questions cannot be answered through the simple repetition of unassimilated information, they have a greater potential for unraveling underlying conceptual structures" (Vosniadou 1994, p. 50). Where possible, we used generative questions written with non-technical language (although some astrophysical terms and jargon were unavoidable).

We also used several well-established techniques of question writing. Each item is limited to a single idea, negatives are avoided, space is provided between multiple parts of a single question, and important words are bolded, italicized, and/or underlined (Henriques, Colburn, and Ritz 2006). These techniques were used in order to clarify what each item is asking.

In the following four sections (Sections 4.3-4.6) I describe in detail the specific items we wrote for each survey. I also defend our selection of these items in the context of the requirements of the design process, as elucidated in this section and Section 4.1 above. Note that the items described in Sections 4.3-4.6 are for the surveys administered during the fall 2009 semester; the surveys were modified for subsequent semesters, as discussed in detail in Sections 6.7 and 7.6.

4.3 Form A Overview

There are five items on Form A, all of which relate to the *Hubble plots* construct. Items 1-4 ask students to select one or more graphs from a bank of eight Hubble plots that correspond to a given situation (constant expansion, constant contraction, accelerating expansion, and decelerating expansion). Item 5 presents students with three Hubble plots, all of which have constant and positive slopes, and asks students to select the plot that corresponds with the oldest universe. All five questions require students to explain their reasoning behind their choices, so simply choosing the correct graph is insufficient to earn a high placement on the construct map in Table 4.1. Since the *Hubble plots* construct focuses on students' abilities to reason about the age and expansion rate of the universe using Hubble plots, we felt that these five items on this survey form adequately covered this construct's domain.

4.4 Form B Overview

Form B has six items on the *models* construct. Three of these items (Items 1, 2, and 6) provide students with opportunities to explain what the expansion of the universe means, what the Big Bang means, and what is expanding in the universe, respectively. The other three items (Items 3-5) epitomize generative questions. Item 3 was meant to determine whether or not students think the Big Bang was an (explosive) event located in empty space. Item 4 addresses whether or not the universe has a center. Item 5 measures if students think there is an edge to the distribution of galaxies in the universe (which, as we found in previous surveys, many students do, especially if they conceive of the Big Bang as "throwing" galaxies out into empty space). We made a deliberate decision on Items 3-5 to not simply ask students if, for example, the universe has a center, because many students can simply regurgitate an answer they have heard in class even if it is not integrated into their fundamental conceptualizations of the expansion of the universe (what Vosniadou 1994 calls "inert knowledge"; see Chapter 2). We believed these questions probe areas of potential difficulty revealed by previous research and cover the full range of attributes listed in Table 4.2.

4.5 Form C Overview

Form C focused on the *evolving universe* construct. It originally included five items, but we threw out item 2 after we realized it contained too much jargon (e.g. elliptical galaxies, active galaxies) and after we realized there was no consensus among astrophysicists about the correct answer. Item 1 asks students how long light will take to travel between two galaxies currently separated by eight billion light-years in an expanding universe. For this question, we asked students to choose their answers from four choices and to explain their selections. We did this in order to simplify the question and possible answer choices for the students while still allowing them to provide us with insights into any potential conceptual or reasoning errors. Items 3-5 ask students if the temperature, total amount of matter, and density of matter have changed over time and to explain their reasonings for their answers. Once again, the items on this survey were chosen based on the information we wanted to learn about students on this construct, as shown in the construct map in Table 4.3.

4.6 Form D Overview

Form D, which corresponds to the *dark matter* construct, has four items, all of which build off students' selections of the rotation curve for a spiral galaxy from six choices in item 1. Item 2 asks students to compare the speeds of three stars at different radii in the galaxy. Item 3 asks students to compare the net gravitational force felt by the three stars; we later threw out this item after we realized we made an incorrect assumption: Namely, stars at greater radii must feel the same or a larger net gravitational force as stars at smaller radii in order for all stars to orbit at the same speed.

the gravitational force is the same whenever the velocities are the same. Item 4 asks students to synthesize their previous answers and infer how matter is distributed throughout the galaxy. Each of these items corresponds to a different attribute on the *dark matter* construct map (Table 4.4).

4.7 External Review

Before we gave these surveys to students, we had multiple groups of education and astrophysics experts review them for clarity, accuracy, and construct coverage. Specifically, our survey items were reviewed by members of the Physics Education Research Group at the University of Colorado, graduate students in the School of Education at the University of Colorado, and astrophysicists at both the University of Colorado and the University of Arizona. In this section, I discuss the modifications we made to the surveys as a result of these reviews.

Ten members of the Physics Education Research Group reviewed a preliminary bank of items we considered for inclusion in the surveys. They did not comment much on the content of the items, since most are not experts in cosmology, but they did help clarify the wording and presentation of several items. We were able to whittle down our pool of items based on their suggestions. We selected a subset of items that survived this review to appear on our first drafts of Forms A-D.

We subsequently gave these drafts to three graduate students in the School of Education. Their input was valuable since they are the only people who reviewed the items who do not identify themselves primarily as practicing physicists or astronomers. These three students examined the questions and pointed out word choices that were potentially confusing for non-experts. For example, item 3 on Form B originally read:

> If you were alive at the time of the Big Bang, would there have been any locations in the early universe from which you could have watched the Big Bang from a

> distance? If yes, describe what it would have looked like and draw a picture. If no, explain why not.

The word “early” did not appear in the version of this item given to Astro 101 students. The three graduate students all agreed that the item made more sense without the word “early.” Including “early” made them over-analyze the response that the question was asking for. Overall, however, they found the wording of the items to be clear.

We revised the surveys based on the recommendations of the three graduate students and then presented the surveys to an expert panel of three astrophysicists drawn from the University of Colorado at Boulder’s Department of Astrophysical and Planetary Sciences. All participating astrophysicists have conducted research in cosmology, taught cosmology in Astro 101, or both. They evaluated the items along three dimensions. First, they looked at whether or not the items would make sense to an Astro 101 student. Second, they considered whether they would expect an Astro 101 student to be able to answer these items post-instruction. Finally, they examined the items for possible errors in the relevant astrophysics. Most of their comments resulted in only minor modifications to the items. For example, the y-axes of the plots in Form A, Items 1-4 originally read “speed.” All members of the panel agreed that “velocity” is the more accurate label. A review of the items by astrophysicists at the University or Arizona had the same result, although they also questioned the premise of item 2 on Form C – an item which we subsequently threw out, as noted above.

At this point, our surveys had been through several rounds of design and development, from defining the constructs to being reviewed by experts. We were now prepared to administer the surveys to Astro 101 students. Appendix A contains copies of the surveys we administered to students in the fall 2009 semester. Chapter 6 analyzes students responses to these surveys.

4.8 Summary of the Survey Design Process

In this chapter, I described the process by which we designed our four conceptual cosmology surveys. This process began by defining the constructs each survey is intended to measure. We then created construct maps for each construct in order to operationalize (to the greatest extent possible) and present in tabular form our hypotheses of how students vary along each construct. At this point, we limited our pool of items to those that could help us place students at appropriate levels of the constructs and that met other constraints (such as limiting the time students need to complete a survey to twenty minutes). I defended the selection of each item for each of the four surveys and discussed the (largely favorable) review of the surveys by education and astrophysical experts. This process is critical for establishing the validity of the surveys, as I discuss in Sections 6.6, 7.5, and 8.5 below.

In Chapter 5 I describe the suite of lecture-tutorials we created and the process by which they were designed. Obviously, the lecture-tutorials address these constructs defined in this chapter and, as such, there are some terms that will reappear in my discussion of the lecture-tutorials. Please keep in mind, however, that there is not necessarily a one-to-one correspondence between a construct and a lecture-tutorial; in some cases, we found that we needed multiple lecture-tutorials to sufficiently cover certain constructs.

Chapter 5

Design of the Lecture-Tutorials

In Chapter 4, I discussed the process by which we designed our four surveys of conceptual cosmology knowledge. This chapter focuses on the design of the cosmology lecture-tutorials. As with many other research-supported curricular materials, the design of the lecture-tutorials lies at the intersection of the purposes they are meant to serve, models of learning pertinent to those purposes, and empirically-tested and theoretically-supported design principles derived from those models (Swan 2008).

A key aspect of the design process elucidated by Schunn (2008) is the need to state the requirements (including, but not limited to, learning goals, audience, and time frame) upfront. Therefore, I begin with a discussion of the constraints imposed by the typical Astro 101 classroom (Section 5.1), followed by a description of our general goals for and their implementation in the cosmology lecture-tutorials (Section 5.2). Section 5.3 provides a brief overview of the content and logic of the suite of five lecture-tutorials we wrote for the fall 2009 semester. Section 5.4 summarizes this chapter.

5.1 Constraints

We designed the cosmology lecture-tutorials for the same Astro 101 environment as the original *Lecture-Tutorials for Introductory Astronomy* (Prather *et al.* 2005; Prather *et al.* 2008). Specifically, we wanted activities that foster the kind of interactive engagement known to promote large learning gains (see Chapter 2) in classrooms with large student-to-instructor ratios, fixed stadium-style seating, and no laboratories or break-out recitation sections. Although many of these characteristics may be non-optimal for student learning, they are nevertheless common to many Astro 101 courses. In the language of Schunn (2008), they represent the constraints over which we must optimize our design of the lecture-tutorials. For example, the large student-to-instructor ratios mean that the cognitive steps between questions on the lecture-tutorials must be small enough that groups of students working together can resolve most difficulties without direct assistance from an instructor (Prather *et al.* 2005). We assume that if the cosmology lecture-tutorials work in these classes, they will work in other environments.

According to the situated perspective of learning (Section 2.3), accounting for the situations in which the lecture-tutorials are used is critical to the success of their implementation. If learning is to be understood as increasing one's ability to participate in culturally-valued activities (Brown, Collins, and Duguid 1989; Greeno 1997; Greeno 2006; Sfard 1998), then the lecture-tutorials must fit within the existing norms and culture of the typical Astro 101 course (Prather *et al.* 2005). As Finkelstein and Pollock (2005) discuss in their study of the implementation of the *Tutorials in Introductory Physics* (McDermott, Shaffer, and the Physics Education Group 2002) at the University of Colorado at Boulder, the physical space in which the tutorials are completed, the training and support of the instructional staff, the level at which students engage with tutorial tasks (e.g. sense making versus "answer getting"), the degree to which the tutorials support the other activities and goals of the course, and the ability of the tutorials to foster productive student-student and student-instructor interactions all determine whether or not the implementation of tutorials will be successful. These points are further emphasized by Black (2009), who stresses that any successful educational materials must affect classroom practices and promote and enrich the nature of feedback interactions that occur between students and instructors. We chose the lecture-tutorial approach to affect student learning precisely because we want resources that the average Astro 101 instructor, who is often busy with research and service requirements in addition to teaching, can integrate into the existing structures of her course with as few modifications to the course as possible (although adopting the lecture-tutorials does require the instructor to reconceptualize her role in the class from lecturer to facilitator; Prather *et al.* 2005).

5.2 Lecture-Tutorial Goals, Implementation Practices, and Design Features

Our primary goal for the cosmology lecture-tutorials is to move students toward more expert-like understandings of the constructs delineated in Chapter 4. We are interested in conceptual, rather than quantitative, understandings. Swan (2008) lists several characteristics of activities that effectively promote conceptual understanding, including (but not limited to)

(1) Using collaborative tasks that force students to communicate and develop their scientific language;

(2) Exposing and discussing common misconceptions and building upon students' prior knowledge; and

(3) Encouraging reasoning over "answer getting."

In this section, I describe the general methods by which we enact these goals given the constraints discussed in the previous section.

Following Prather *et al.* (2005), each lecture-tutorial is a two to six page worksheet on a single topic, composed of Socratic-style questions that students are expected to complete in groups of two or three in an average of fifteen minutes. These worksheets are designed for use during lecture; in the traditional implementation, students work on a lecture-tutorial after receiving a brief lecture that introduces them to the basic terms and concepts they need in order to make sense of the activity (Prather *et al.* 2005). The lecture-tutorials avoid jargon in favor of students' natural language whenever possible in order to improve student comprehension. Furthermore, we followed some of the recommendations of Siegel, Wissehr, and Halverson (2008) to make the lecture-tutorials accessible to students who are learning English as a second language. Instructors frequently bookend a lecture-tutorial with a series of conceptually challenging think-pair-share questions and often leave a few minutes post-lecture-tutorial for a debriefing session in order to clarify any residual confusion

(Prather *et al.* 2005). These design features and implementation practices address Swan's (2008) first requirement in the above paragraph.

Swan's (2008) second requirement is addressed by the use of conceptual change stategies such as elicit-confront-resolve (Heron 2004b; McDermott 1991) and bridging (Brown and Clement 1989; Clement, Brown, and Zeitsman 1989; Elby 2001) throughout the lecture-tutorials. Thus, our design of the lecture-tutorials is significantly guided by cognitive models of learning (Section 2.2). Whenever possible, we account for the knowledge students bring with them to the classroom – which means the lecture-tutorials must not depend on ideas that are likely to be underdeveloped in or foreign to novices (Heron 2004b). As noted by Heron (2004b), with regards to the *Tutorials in Introductory Physics*, many of our activities may parallel chains of reasoning found in many textbooks. The difference between a textbook and the lecture-tutorials is that the lecture-tutorials help guide the students through the process of constructing the chains of reasoning for themselves. Ultimately, we share with Heron (2004b) the following goal: "Students need to learn to construct, articulate, criticize, and defend logical arguments and to recognize the proper relationship between intuition and reason if an understanding of the underlying concepts is to be meaningful" (p. 364). Conceptual change strategies suggested by cognitive models of learning have previously been shown to help Astro 101 students achieve these goals (Prather *et al.* 2005; Hudgins *et al.* 2006; LoPresto and Murrell 2009).

Finally, how do the cosmology lecture-tutorials promote reasoning over "answer getting?" First, the questions in each lecture-tutorial prompt students to explain their reasoning throughout the activity. This is complemented by certain implementation practices adopted by the instructor, such as allowing students sufficient time to explain their reasonings (typically five to eight minutes per page, on average), requiring that groups reach a consensus on the answer to each question before proceeding, and following the lecture-tutorial with conceptually challenging think-pair-share questions that require students to have improved their conceptual understandings during the lecture-tutorial. Second, the tutorials require students to interpret data tables, figures, and graphs. Developing a fluency with multiple representations has been shown to be important in developing students' conceptual understandings and promoting more expert-like approaches to a topic (Kohl and Finkelstein 2006; Kohl, Rosengrant, and Finkelstein 2007; Kohl and Finkelstein 2008; Van Heuvelen 1991). Finally, each lecture-tutorial contains one or more "student debates," in which two or more fictional students argue about the answer to a prior question. One of these students typically expresses a common reasoning or conceptual error we have observed with real Astro 101 students, while a different student presents a more accurate answer (but always in Astro 101 students' natural language). These student debates are designed as a way to "catch" students or groups of students that have progressed through the lecture-tutorial with their incorrect ideas intact; the student debates are meant to force them to reconsider their conceptual understandings of the material and may help promote cognitive change (Prather *et al.* 2005) and metacognitive skills (Schoenfeld 1987). These approaches work in concert to emphasize the importance of reasoning over simply obtaining an answer.

This section detailed the general principles and goals that guided our writing of the cosmology lecture-tutorials. In the following section, I describe how we implemented these principles in each of the lecture-tutorials.

5.3 Overview of the Cosmology Lecture-Tutorials

We authored a suite of five cosmology lecture-tutorials. The lecture-tutorials are named "Hubble's Law"; "Making Sense of the Universe and Expansion"; "Expansion, Lookback Times, and Distances"; "The Big Bang"; and "Dark Matter". I briefly describe each of these lecture-tutorials in the following subsections and discuss how they enact conceptual change strategies. See Appendix B for full copies of these lecture-tutorials.

5.3.1 Hubble's Law

The "Hubble's Law" lecture-tutorial focuses on helping students understand Hubble's law and interpreting Hubble plots. It begins with a drawing of four galaxies with the (proper) distances between each labeled. We ask students to redraw this picture after the universe has doubled in size.

We found that most students intuitively and correctly double the distances between the galaxies. From this intuitive starting point, we then lead students through a series of questions (including a student debate) that build upon this result to help them understand why, in a universe that is expanding at a constant rate over time, farther galaxies appear to move away from us at higher velocities and why the Hubble plot for this universe is a straight line with a positive slope. We then ask students to consider how the Hubble plot would change if the universe had expanded at a faster or slower (but constant) rate and what affect this would have on the age of the universe. Finally, we ask students to interpret a Hubble plot for an accelerating universe and describe what this means in terms of our ability to see galaxies in the distant future.

5.3.2 Making Sense of the Universe and Expansion

This lecture-tutorial addresses the common misconception that an expanding universe must have a center and an edge. The tutorial begins with a drawing of a circle filled with galaxies, which represents our observable universe. The position of the Milky Way at the center of this circle is marked, as is the location of "Galaxy X," which resides at the very edge of our observable universe. We ask students whether or not there are more galaxies beyond our observable universe; most agree that there should be. We then ask students to draw a circle representing Galaxy X's observable universe. Students then encounter a student debate in which one student argues there must be more galaxies beyond Galaxy X, while the other claims that Galaxy X is simply the last galaxy before an enormous void of nothingness. By this point, we hope that most students accept the idea that the universe is filled with galaxies and that, on the largest scales, the universe looks the same in every direction one looks and from every location. We have thus introduced students to the cosmological principle without ever referring to it as such. We then ask students whether the universe has an edge or a center. Most say no, based on their previous answer. This sequence of questions exemplifies the bridging/refining intuitions strategy (Brown and Clement 1989; Clement, Brown, and Zeitsman 1989; Elby 2001) since it starts with an idea most students agree with and builds to a conclusion many disagree with coming in to Astro 101.

In the second part of the lecture-tutorial, we present students with the balloon analogy for the expansion of the universe. We have two purposes for doing this. First, we want to provide students with a model of how something can expand and yet lack a center and an edge. We thus have several questions that ask students if they ever encounter a center or an edge as they travel along the skin of the balloon and whether or not one sees galaxies moving away from one's location irrespective of where one is on the balloon. Second, we use the balloon analogy to stress the limits of models and analogies in science. We thus ask students a series of questions that force them to state whether or not a specific aspect of the real universe is accurately portrayed by the balloon analogy. We have found that students who do not go through this exercise often maintain deep confusions about the nature of expansion (see Chapters 6-8).

5.3.3 Expansion, Lookback Times, and Distances

This lecture-tutorial qualitatively addresses the relationship between light travel times, lookback times, and distances in an expanding universe. I discussed this issues quantitatively in Section 3.3. Students are first presented with a picture of two galaxies (Galaxy A and Galaxy B) and told their (proper) distance from one another (3 billion light-years) and the age of the universe (4 billion years) when a star explodes in Galaxy B. We ask students how long the light from the explosion takes to travel between Galaxies A and B, how far the light traveled, the age of the universe at the end of the light's journey, and the lookback time associated with that light in the context of a static (i.e. non-expanding) universe. We observe that most students can answer these questions and they provide an important starting point for the next set of questions. The next set of questions are similar, except we now ask students to qualitatively consider the effects of expansion. For example, question 6 asks students the following:

> By the time the light from the explosion reaches Galaxy A, is the distance to Galaxy B more than, less than, or exactly 3 billion light-years?

The lecture-tutorial ends with a question that asks students to summarize how expansion affects lookback times and light-travel times.

5.3.4 The Big Bang

The "Big Bang" lecture-tutorial explicitly confronts the erroneous idea that the Big Bang was an explosion of pre-existing matter into empty space. The lecture-tutorial begins with two drawings: One shows a series of dots enclosed within a square that get farther from one another as the square expands. The other shows the same group of dots getting farther from one another as they move into the empty space of a square of constant size. We ask students which drawing represents expansion and which represents an explosion. We then present students with a student debate before asking them which is the better representation of the history of the universe. Most students select the expanding square.

We next present students with a series of questions on how the temperature and density of the universe change over time. We ask students to describe how the universe would change if its history played like a movie running backward. The lecture-tutorial ends with a student debate and a two questions that lead students to the idea that the Big Bang Theory refers to the expansion of the universe from an initially hot and dense state and that matter formed from energy via Einstein's famous equation $E = mc^2$.

Note that in this lecture-tutorial we referred to the early universe as being small in size (for example, in questions 12 and 13). As mentioned in Section 3.2, this is not necessarily an accurate statement. We corrected this error in later versions of this lecture-tutorial.

5.3.5 Dark Matter

The goal of the "Dark Matter" lecture-tutorial is to enable students to explain why flat rotation curves in spiral galaxies are evidence for dark matter. The lecture-tutorial begins by telling students that an object's orbital velocity depends on the mass inside its orbit. We then present students with a table listing the planets of the Solar System, their orbital speeds, and the mass inside their orbits. Students then answer a series of questions based on this table to help them realize that most of the mass in the Solar System must be located in the Sun and, consequently, the orbital velocities of planets drops with increasing distance from the Sun.

In the second part of the lecture-tutorial, we show students a drawing of the Milky Way Galaxy and ask them where most of the mass appears to be located. Most students claim that the mass is concentrated in the center of the galaxy. We then ask a series of questions that *elicits* the idea that the rotation curve for the Milky Way should look like the rotation curve for the Solar System. We then *confront* students with the true rotation curve for the Milky Way and ask a series of questions to help students *resolve* the discrepancy between their expectations and reality. The lecture-tutorial ends with multiple questions that require students to explicitly state that there must be more mass in the Milky Way than we can detect with light.

Note that just like Form D (as described in Section 4.6), we made the mistake in this lecture-tutorial of assuming that stars at greater radii must feel the same or a larger net gravitational force as stars at smaller radii in order for all stars to orbit at the same speed.

5.4 Summary of the Lecture-Tutorial Design Process

In this chapter, I detailed the principles and constraints that guided the development of the cosmology lecture-tutorials. While the lecture-tutorials themselves are heavily influenced by cognitive models of conceptual change, I also briefly described how the situated perspective of learning influenced our design choices, especially in light of the constraints of the typical Astro 101 class. Our goal was to develop materials that will be readily incorporated into existing Astro 101 courses and help students develop more expert-like understandings of important cosmological topics.

Of course, the design of many effective educational materials is often an iterative process (Schunn 2008). One should expect revisions based on students' responses to the materials and the assessments used to judge the efficacy of the materials (Heron 2004b). The cosmology lecture-tutorials described in Section 5.3 above are our initial drafts for the fall 2009 semester. We modified these lecture-tutorials for subsequent semesters as we collected more data. These modifications are described in more detail in the following chapters.

Chapter 6

Fall 2009 Results

This is the first of three chapters presenting the results of our study. This chapter looks at the data we collected in the fall 2009 semester, while Chapters 7 and 8 examine the data from the spring and fall 2010 semesters, respectively. We made a deliberate choice to breakdown our results by semester and present them chronologically in order to give the reader a feel for the evolutionary approach we took: After each semester, we revised our surveys and lecture-tutorials based on our findings. Taken together, Chapters 6-8 show how we refined our instruments based on the data we collected over the past three semesters.

Throughout this and the following two chapters, we try to answer the two major questions of this project:

(1) What are the common conceptual and reasoning difficulties Astro 101 students encounter when studying cosmology?

(2) Do the cosmology lecture-tutorials help students overcome these difficulties? (Or, phrased another way, do students who use the cosmology lecture-tutorials exhibit larger learning gains than their peers who do not?)

In order to help answer these questions, we also examine the data to help us judge the efficacy of our primary measurement instruments: the four conceptual cosmology survey forms.

This chapter has several major components. Section 6.1 describes the classes that participated in this study in the fall of 2009. Section 6.2 briefly explains the scoring rubrics we developed for each item on each survey. Much of this chapter focuses on applying two psychometric measurement models, classical test theory (CTT) and item response theory (IRT), to the data; these models and their results are described in more detail in Sections 6.3 and 6.4. Section 6.5 examines students written responses in detail and highlights the common difficulties of Astro 101 students. I then consider whether or not we have evidence to support the validity of our surveys – that is, can we claim they are measuring what we intended them to measure (see Section 6.6)? Finally, in Section 6.7, I discuss the revisions we made to the surveys and the lecture-tutorials as a result of the findings presented in this chapter. Section 6.8 summarizes this chapter.

6.1 Surveyed Classes

Three classes participated in this study. All were large lecture-based Astro 101 courses. Class A was taught at the University of Arizona and use the cosmology lecture-tutorials in class, following the implementation recommendations in Prather *et al.* (2005). Class B was taught at the University of Colorado at Boulder and used the lecture-tutorials in smaller (20-25 student) recitation sections. All students enrolled in this class were required to attend their weekly recitation sections. The recitations were typically run by undergraduate learning assistants (see Otero *et al.* 2006 and Otero *et al.* 2010 for more information about the learning assistant program); these learning assistants met weekly with the course instructor to prepare and coordinate their plans for the upcoming week's recitation. Class C was also taught at the University of Colorado at Boulder. It did not have any recitation, laboratory, or other class meeting times outside of its lectures three times per week. Class C was our control group as it did not use the lecture-tutorials. We originally recruited a second Astro 101 class for this semester to participate in the study and not use the lecture-tutorials. Unfortunately, we had to eliminate this class from the study after we learned that the instructor covered some cosmology material before we could administer the surveys pre-instruction. Thus, Class A, Class B, and Class C were the only participating courses for the fall 2009 semester. The number of students participating from each class is shown in Table 6.1.

We administered the four survey forms both before and after instruction. All four surveys

Table 6.1: Number of participants pre- and post-instruction per class for fall 2009.

Class	Pre-Instruction	Post-Instruction
Class A	282	231
Class B	119	100
Class C	100	75

were administered once and each student only responded to a single survey. This means each survey was given to approximately a quarter of each class during each administration. We distributed the surveys randomly among the students each time we surveyed a class, so we made no effort to match pre- and post-test data. In the discussion of our results below, we focus strictly on the pre- and post-test averages.

6.2 Scoring Rubrics

We constructed detailed scoring rubrics for each item on each survey. These rubrics allow us to score and categorize the full range of students' responses – which is essentially Wilson's (2005) third step in survey design and interpretation (see Chapter 4 for a list of all four steps). We constructed our rubrics only when we had all the pre- and post-test responses in hand. Our rubrics are therefore based on a detailed, iterative, qualitative analysis of actual student responses. For most items, the rubric has two components: An overall score, which is based on whether or not the student gives a correct answer and a complete and correct explanation, and codes for the most common reasoning elements used by students. These rubrics allow us to perform quantitative analyses on the survey responses (via the overall scores - see Sections 6.3 and 6.4 below) as well as examine the data for patterns in which reasoning elements are used by students pre- and post-instruction, with or without the lecture-tutorials (see Section 6.5). See Appendix C for the scoring rubrics for each survey form for the fall 2009 semester.

6.3 Classical Test Theory Analysis

Wilson's (2005) fourth step in survey design and interpretation is applying psychometric measurement models to the data. This allows us to actually answer the question of whether or not the lecture-tutorial students outperform the non-lecture-tutorial students. They also help us judge the efficacy of the four survey forms. We used two measurement models in this project: Classical Test Theory (CTT), which is the focus of this section, and Item Response Theory (IRT), which is addressed in Section 6.4.

6.3.1 Classical Test Theory Overview

CTT postulates that a student p's observed score (X_p) differs from her true score (T_p) by a certain amount of error (E_p),

$$X_p = T_p + E_p \tag{6.1}$$

(Lord and Novick 1968). The true score T_p is defined as the expectation value of all the observed scores the student would have earned if she was brainwashed (so as to forget what answers she gave) and took the test multiple times under the same conditions (Lord and Novick 1968). From this simple model, a number of elegant statistics are derived to assess the quality of a test.

One such class of statistics are those used to estimate the *reliability* of the test. A test's reliability can be conceptualized as the ratio of the variance in true scores to the variance in observed scores (Lord and Novick 1968; Wainer and Thissen 2001). When this ratio is close to zero, variations in observed scores from test taker to test taker are mostly due to error and the test is considered to function poorly; conversely, a ratio close to one indicates that the test is working, since much of the observed score variance is due to true score variance (Lord and Novick 1968; Wainer and Thissen 2001).

Unfortunately, the reliability of a test cannot be directly calculated since true scores are, for practical purposes, unobservable. This is why psychometricians have invented several techniques to estimate the reliability of a test. Most estimation procedures involve calculating the correlation between two parallel tests or between two parallel halves of the same tests (Wainer and Thissen 2001). In order for two forms or halves of a test to be parallel, they must have equal true scores and error variances (Lord and Novick 1968). The most popular estimate of reliability is Cronbach's α, which is the average of all possible split-half correlations for a test (Lord and Novick 1968; Wainer and Thissen 2001). Cronbach's α can be written as

$$\alpha = \frac{N}{N-1} \frac{\sigma_x^2 - \sum\limits_{i=1}^{N} \sigma_{y_i}^2}{\sigma_x^2} \ , \tag{6.2}$$

where N is the number of test items, σ_x^2 is the total test score variance, and $\sigma_{y_i}^2$ is the variance of item y_i (Thompson 2003). Note that in the numerator of Equation (6.2) one subtracts the variances of all the items from the total test score variance. This leaves only the covariances among the items. When these covariances are high, Cronbach's α is close to unity and the test is considered to be "internally consistent" (Thompson 2003). Likewise, Cronbach's α approaches zero as the covariances among items decrease (Thompson 2003). To maximize the reliability of a test, Cronbach's α should be as close to one as possible (Wainer and Thissen 2001), although values greater than 0.70 are generally considered acceptable (George and Mallory 2009). Note that Cronbach's α is always a lower bound on the reliability of a test (Lord and Novick 1968).

While Cronbach's α and other estimates of reliability are measures of the overall quality of a test, CTT also has statistics that apply to individual items. For example, the *difficulty* of an item (or item P-value) is defined as the average score on that item divided by the total number of possible points for the item (Lord and Novick 1968). This means that harder items have smaller P-values. The *discrimination* of an item is frequently measured by the correlation between the item scores and total test scores; when this is done for a dichotomously scored item, this correlation is called the point-biserial of the item (Lord and Novick 1968). Survey developers frequently assess the quality of their instruments by looking at the difficulties and discriminations of the survey items. For example, some studies may accept items with P-values as low as 0.10 or 0.20 and as high as 0.80 or 0.90 (Bardar *et al.* 2007; Ding *et al.* 2006; Maloney *et al.* 2001), although P-values should lie close to 0.50 in order to maximize test reliability (Ding and Beichner 2009). Likewise, convention

Table 6.2: The discriminations of the items on Forms A-D for fall 2009.

Form A		Form B		Form C		Form D	
Item	Discrimin.	Item	Discrimin.	Item	Discrimin.	Item	Discrimin.
Item 1	0.63	Item 1	0.70	Item 1	0.68	Item 1	0.58
Item 2	0.63	Item 2	0.65	Item 3	0.81	Item 2	0.59
Item 3	0.63	Item 3	0.58	Item 4	0.61	Item 4	0.82
Item 4	0.58	Item 4	0.63	Item 5	0.75		
Item 5	0.55	Item 5	0.51				
		Item 6	0.39				

suggests that point-biserials should be greater than or equal to 0.20 (Ding and Beichnner 2009), although researchers are free to set their own criteria (e.g. Bardar *et al.* 2007 sought items with point-biserials between 0.30 and 0.70 for the *Light and Spectroscopy Concept Inventory*). In Section 6.3.2 we look at both the P-values and item-test score correleations as measures of the difficulties and discriminatory powers, respectively, of the items on our surveys.

6.3.2 Items' Difficulties and Discriminations

The discriminations of the items on Forms A-D for the fall 2009 semester are shown in Table 6.2. All fall above the conventionally accepted minimum values cited in Section 6.3.1 and most are relatively high (> 0.50). The one exception is Item 6 on Form B. I will discuss one hypothesis for this item's low discrimination in Section 6.5.2 below.

Table 6.3 shows the P-values for the items on Forms A-D for the fall 2009 semester. CTT statistics, including P-values, are highly sample-dependent (Hambleton and Jones 1993; Thompson 2003). P-values, for instance, necessarily depend on how much students know about the construct being measured. As students learn more about a construct, the item's P-values must change (Wallace and Bailey 2010). Therefore, in addition to the overall P-values (which we calculated using every pre- and post-instruction response), we also calculated the pre- and post-instruction P-values for the students who used the lecture-tutorials ("LT pre" and "LT post") and their peers who did not ("Non-LT pre" and "Non-LT post").

The different P-values for these two populations of students are our first pieces of evidence

Table 6.3: The difficulties (P-values) of the items on Forms A-D for fall 2009.

	Item	Overall	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Form A	Item 1	0.48	0.45	0.50	0.44	0.57
	Item 2	0.39	0.39	0.40	0.37	0.43
	Item 3	0.36	0.34	0.40	0.33	0.33
	Item 4	0.35	0.34	0.38	0.33	0.33
	Item 5	0.50	0.52	0.47	0.58	0.48
Form B	Item 1	0.50	0.36	0.63	0.53	0.68
	Item 2	0.66	0.52	0.81	0.68	0.69
	Item 3	0.70	0.61	0.79	0.74	0.71
	Item 4	0.56	0.41	0.72	0.56	0.60
	Item 5	0.75	0.71	0.86	0.59	0.74
	Item 6	0.66	0.61	0.70	0.66	0.75
Form C	Item 1	0.67	0.51	0.85	0.69	0.68
	Item 3	0.59	0.39	0.83	0.50	0.69
	Item 4	0.51	0.41	0.67	0.41	0.46
	Item 5	0.57	0.43	0.75	0.57	0.50
Form D	Item 1	0.68	0.57	0.89	0.53	0.61
	Item 2	0.89	0.86	0.90	0.91	0.92
	Item 4	0.63	0.55	0.74	0.57	0.68

that the cosmology lecture-tutorials are having positive effects on students' performances. For most of the items in Table 6.3, the average score increases by a greater amount for students who used the lecture-tutorials than for students who did not. The only exceptions to this pattern are the aforementioned Item 6 on Form B as well as Items 1, 2, and 5 on Form A. I will discuss these items in more detail below.

6.3.3 Reliability

Table 6.4 shows Cronbach's α for Forms A-D for the fall 2009 semester. All these values are lower than conventionally accepted minimum values (although Cronbach's α for Form C is close to the recommended $\alpha = 0.70$ of George and Mallory 2009). Why is Cronbach's α for each form so low?

Like many CTT statistics, Cronbach's α depends on the test-taking population and the items to which they respond. For example, homogeneous populations will yield lower values of Cronbach's α than heterogeneous populations, since Cronbach's α depends on the total score vari-

Table 6.4: Cronbach's α for Forms A-D for fall 2009.

Form	Cronbach's α
Form A	0.52
Form B	0.60
Form C	0.68
Form D	0.40

ance (Thompson 2003). As I describe below, this may be a factor in the value we calculate for Form A. Furthermore, Cronbach's α is also sensitive to the brevity of a test (Schmitt 1996). Shorter tests typically yield smaller values. One can see how this estimate of a test's reliability grows with an increasing number of items via the Spearman-Brown prophesy formula (Wainer and Thissen 2001). This formula relates the old estimate of reliability α_{old} to the improved estimate α_{new} that one would achieve if there were M times as many items on the test:

$$\alpha_{new} = \frac{M\alpha_{old}}{1 + (M-1)\alpha_{old}} \tag{6.3}$$

(Lord and Novick 1968). Figure 6.1 shows how the reliability of Forms A-D would improve if we increased the number of items on each test by a factor M. Since these conceptual cosmology surveys are short (especially after we threw out Item 2 on Form C and Item 3 on Form D), their values of Cronbach's α are smaller than they would be if we could have used longer surveys (which would conflict with the administration time requirements noted in Section 4.2).

Another form of reliability we investigated, in addition to Cronbach's α, is called inter-rater reliability. The inter-rater reliability component of our study looked at whether or not other science education researchers can use the scoring rubrics to arrive at the same overall scores and assign the same student response codes to a subset of items and students (Otero and Harlow 2009). Specifically, two science education researchers collaborated to score 65 responses. These responses were taken from 21 students and 9 items. We then compared their scores to those I assigned. We agreed on 83% of the item scores. Furthermore, 74% of the codes I assigned to students' responses agreed with the codes the other researchers assigned. Likewise, 76% of the codes the other researchers assigned agreed with the codes I assigned. These results support the inter-rater

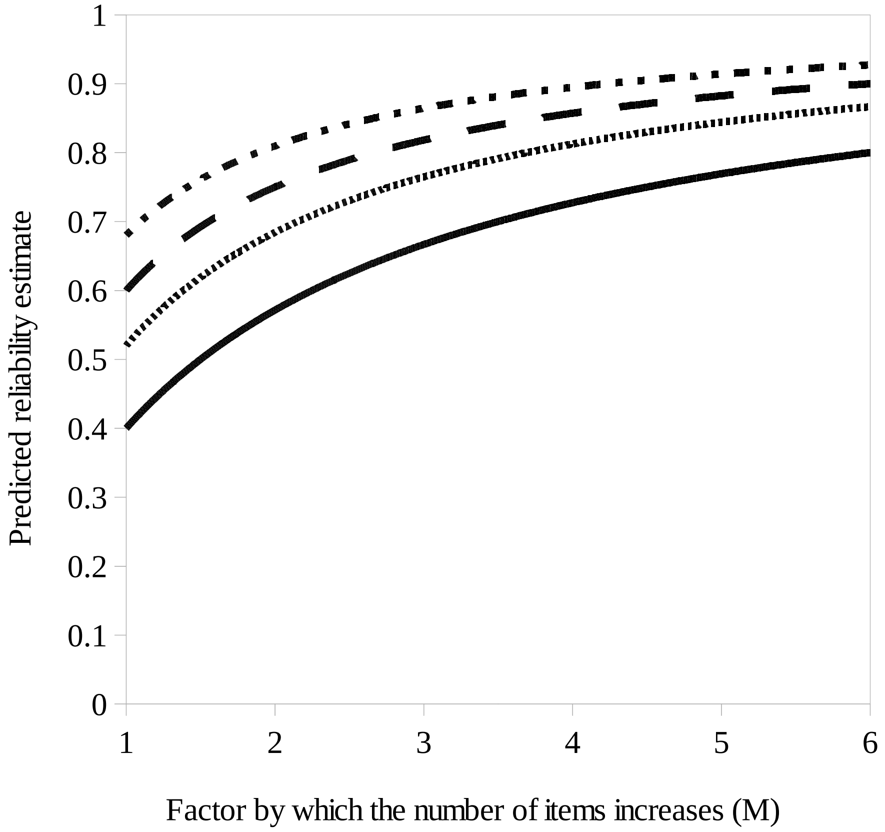


Figure 6.1: How the estimated reliabilities of Forms A-D would change if the number of their items increased by a factor M. Form A is the dotted line, Form B is the dashed line, Form C is the dashed and dotted line, and Form D is the solid line.

reliability of our results.

Cohen's κ (Cohen 1968) provides another way to ascertain inter-rater reliability. It may be defined as

$$\kappa \equiv \frac{q_o - q_e}{1 - q_e} \ . \tag{6.4}$$

In Equation (6.4), q_o represents the overall proportion of scores on which two raters agree (Fleiss, Levin, and Paik 2003). The quantity q_e represents the overall proportion of scores on which the two raters are expected to agree by chance (Fleiss, Levin, and Paik 2003). One can calculate q_e from the observed data: For each score category i, multiply the proportion of cases one rater assigns i by the proportion of cases by which the other rater assigns i. If there are a total of k score categories, this leaves one with k numbers. Summing these k numbers yields q_e. Cohen's $\kappa = 0.755$

for our data. Fleiss, Levin, and Paik, (2003) say such a value for κ corresponds to an "excellent" agreement, while Landis and Koch (1977) call it a "substantial" agreement.

One weakness of Cohen's κ is that it equally weights all deviations between raters' scores (Cohen 1968; Fleiss, Levin, and Paik 2003). For example, an item for which I assigned a student a score of 3 and another rater assigned a score of 1 receives the same weight in the calculation of Cohen's κ as an item for which I assigned a student a score of 3 and another rater assigned a score of 2. Cohen (1968) introduced a modification of his eponymous statistic in order to correct this weakness. Cohen's weighted κ (κ_w) is defined as

$$\kappa \equiv \frac{q_{o(w)} - q_{e(w)}}{1 - q_{e(w)}} \ , \tag{6.5}$$

where $q_{o(w)}$ is the observed weighted proportion of agreement and $q_{e(w)}$ is the chance-expected weighted proportion of agreement (Fleiss, Levin, and Paik 2003). The observed weighted proportion of agreement $q_{o(w)}$ is given by

$$q_{o(w)} = \sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij} q_{ij} \ , \tag{6.6}$$

where q_{ij} is the proportion of cases for which one rater assigns a score of i and another assigns a score j and w_{ij} represents the weight associated with q_{ij}. The chance-expected weighted proportion of agreement $q_{o(w)}$ given by

$$q_{e(w)} = \sum_{i=0}^{k}\sum_{j=0}^{k} w_{ij} q_{i.} q_{.j} \ , \tag{6.7}$$

where $q_{i.}$ is the proportion of cases one rater assigns i and $q_{.j}$ is the proportion of cases the other rater assigns j (Fleiss, Levin, and Paik 2003). There are various ways to define the weights; for our purposes, we used the linear weights given by

$$w_{ij} = 1 - \frac{|i-j|}{k-1} \ . \tag{6.8}$$

For our data, $\kappa_w = 0.823$. This again corresponds to what Fleiss, Levin, and Paik (2003) call an "excellent" agreement. Landis and Koch (1977) consider this to be an "almost perfect" agreement. Regardless of whether one adopts the characterizations of Fleiss, Levin, and Paik (2003) or Landis and Koch (1977), and regardless of whether one uses κ or κ_w, our inter-rater reliability appears to be strong.

6.3.4 Normalized Gains

How do the students who used the cosmology lecture-tutorials compare to the students who did not? Table 6.3 shows a greater, positive change in the P-values of many items for students who used the lecture-tutorials compared to students who did not. This indicates that, for many items, students who used the lecture-tutorials outperform their peers who did not. However, Table 6.3 also shows that, for some items, the lecture-tutorial population of students started from a different place than the non-lecture-tutorial population. This is further supported by the data presented in Section 6.5. The fact that the two populations often have different starting points suggests that a more equitable comparison should use Hake's normalized gain (Equation (2.1)).

Figure 6.2 shows the normalized gains for each item on each form, as well as for the overall scores on each form. With the exception of Item 6 on Form B, the normalized gains for the students using the lecture-tutorials are always higher than the normalized gains of students who did not on Forms B-D. The normalized gains on Form A are a different story: Both populations of students exhibit low gains across the board. In some cases, the gains were actually negative. This issue is addressed in more detail in later in this chapter.

An important issue we considered is whether or not these gains are statistically significant. This can actually be broken up into two questions. First, for a given population (lecture-tutorial or non-lecture-tutorial), are the differences in the pre- and post-instruction scores different from one another at a statistically significant level? Second, are the gains observed for one population different, at a statistically significant level, than the gains observed in the other population? We can answer the former question now; the latter must wait until Section 6.4.6.

Why can we not answer the second question now? Because we do not have matched pre- and post-instruction data, which means we do not have any way to measure the variances or uncertainties in our normalized gains. We will, however, be able address this question in Section

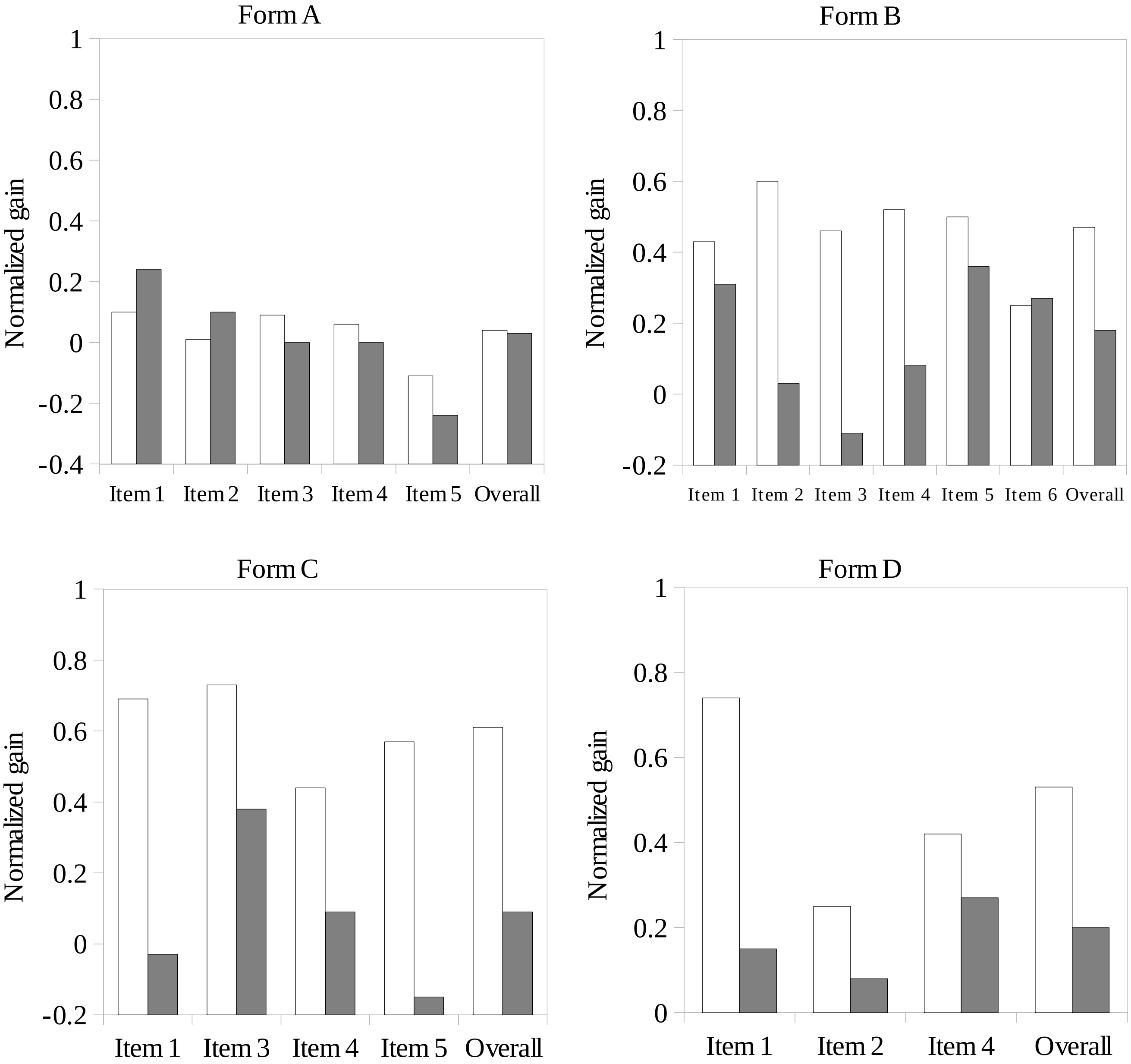


Figure 6.2: The normalized gains for Forms A-D for fall 2009. White bars correspond to the lecture-tutorial population and grey bars correspond to the non-lecture-tutorial population.

6.4.6 because item response theory assigns a standard error of measurement to each estimated student ability; we can thus find and compare the average abilities of each population by using these standard errors of measurement.

Whether or not the differences in the pre- and post-instruction scores for each population are statistically significant can be answered immediately using the Mann-Whitney test. The Mann-Whitney test compares two sets of independent ordinal data (Guilford and Fruchter 1978; Wilcox 1987). “Ordinal” refers to the fact that the numbers in the data set represent an empirical ordering between the objects they represent (e.g. a student with a higher score knows more about a particular construct than a student with a lower score), but the differences between and ratios of these numbers are meaningless (Stevens 1946). Wright and Linacre (1989) argue that raw test scores are typically ordinal. Wallace and Bailey (2010) likewise argue that gain scores are ordinal. Given the likely ordinal nature of our data, the Mann-Whitney test is the appropriate significance test to apply.

The procedure for applying the Mann-Whitney test begins by combining all the data from two data sets (in our case, the pre- and post-instruction raw scores on a survey for either the lecture-tutorial or non-lecture-tutorial students) and rank ordering them (Guilford and Fruchter 1978; Wilcox 1987). The null hypothesis for this test is that the distribution of ranks is the same for both groups (Guilford and Fruchter 1978; Wilcox 1987). In other words, the probability that a given data point from the pre-instruction group has a rank that exceeds the rank of a data point from the post-instruction group is just as probable as the rank of a data point from the post-instruction group exceeding the rank of a data point from the pre-instruction group (Wilcox 1987). The alternative hypothesis is that the higher ranks are preferentially found in the post-instruction group.

In a traditional application of the Mann-Whitney test, one calculates a statistic called U. For our study, U is found by looking at each data point in the pre-instruction data set, counting the number post-instruction data points with smaller ranks, and summing these counts. The distribution of U is is known and can be looked up. However, for samples suitably large (at least eight entries in each data set), the distribution of U is approximately Gaussian with a known mean and standard deviation (Guilford and Fruchter 1978). Since our data is well within these Gaussian limits, we check the significance of U by using the Gaussian distribution to calculate the one-tailed probability of obtaining this value of U. Table 6.5 shows these Mann-Whitney p-values (not to be confused with the item P-values, which are measures of item difficulty) for both the lecture-tutorial

Table 6.5: Mann-Whitney p-values for the lecture-tutorial (LT) and non-lecture-tutorial (Non-LT) groups for Forms A-D for fall 2009.

Population	Form A	Form B	Form C	Form D
LT	0.1003	< 0.0001	< 0.0001	< 0.0001
Non-LT	0.2514	0.0838	0.1922	0.1230

and non-lecture-tutorial populations. As Table 6.5 shows, the gains are not statistically significant ($p < 0.05$) for any form for the non-lecture-tutorial class. They are statistically significant for Forms B-D for the lecture-tutorial population.

6.4 Item Response Theory Analysis

The previous section examined our data from the perspective of classical test theory. Although CTT has several advantages (e.g. its assumptions are weak and easily met, and its statistics are easily computed; Hambleton and Jones 1993), it also has several disadvantages. These include the fact that its statistics are sample-dependent and that the fundamental assumption of CTT (observed scores are the true scores plus some error) cannot be falsified (Hambleton and Jones 1993). Item response theory (IRT) is a complementary psychometric model that does not possess these weaknesses. In this section, we use IRT to analyze our data.

6.4.1 Item Response Theory Overview

The simplest IRT model is the Rasch model (Rasch 1960). It can be written as

$$P(X_{pi} = 1|\theta_p, b_i) = \frac{\exp[\theta_p - b_i]}{1 + \exp[\theta_p - b_i]} \ . \tag{6.9}$$

Equation (6.9) represents the probability that a person p correctly answers a dichotomously scored item i. X_{pi} is the person's response to the item; it is one when the person gives the correct answer and zero when the person gives an incorrect answer. There are two parameters that influence a person's probability of success: the person's ability θ_p and the item's difficulty b_i. These parameters are considered innate properties of the person and item, respectively. A person's ability does not depend on the particular items to which she responds, and an item's difficulty does not depend on the population of people responding to it (Rasch 1960). Note that terms such as "item difficulty" also appear in CTT but have different meanings in the context of IRT.

Equation (6.9) may be rewritten as

$$\ln \left[\frac{P(X_{pi} = 1)}{P(X_{pi} = 0)} \right] = \theta_p - b_i \ . \tag{6.10}$$

This equation shows that in the Rasch model person abilities and item difficulties are measured on the same scale. Their difference equals the natural logarithm of the odds (probability of success divided by probability of failure) of a correct response. This relationship also shows why abilities and difficulties are measured in log odds units (logits). The difference in logits between θ_p and b_i conveys information about the probability that person p correctly answers item i. For example, when $\theta_p = b_i$ the probability of a correct answer is 50% (in fact, one way to define the difficulty of an item in the Rasch model is by the ability one must possess in order to have a 50% chance of giving the correct answer). For more information on the derivation, interpretation, and estimation of the parameters of the Rasch and other IRT models, see the foundational works of Lord and Novick (1968) and Rasch (1960), the pedagogical treatments of Embretson and Reise (2000), Hambleton and Jones (1993), Harris (1989), and Wallace and Bailey (2010), and the references therein.

The Rasch model makes two fundamental assumptions. First, it assumes the test is *unidimensional* – that is, it only measures abilities on a single construct (Embretson and Reise 2000). Second, it assumes *local independence*. Local independence means that the model parameters should explain all correlations between examinees' responses (Embretson and Reise 2000; Yen 1993). Form D manifestly violates the assumption of local independence since it exhibits what Yen (1993) calls "item chaining." Item chaining means that each item builds off the previous item such that knowing the answer to one item increases the probability that one correctly answers the next. Because Form D, by construction, violates the assumption of local independence, we will not analyze it using item response theory.

Other IRT models add additional parameters and/or allow one to relax some of the assumptions listed above. For example, multi-parameter IRT models add parameters that allow items to have different discriminations and account for respondent guessing (Embretson and Reise 2000; Hambleton and Jones 1993; Harris 1989; Lord and Novick 1968). Multidimensional IRT models relax the assumption of unidimensionality (Ackerman, Gierl, and Walker 2003; Briggs and Wilson 2003). For our purposes, however, we are interested in an extension of the Rasch model known as the partial credit model (Masters 1982).

The partial credit model does not assign a single number to an item to represent its difficulty. Instead, it characterizes an item by a collection of step difficulties b_{jk} which determine when a student of ability θ_p is just as probable to have a score j as she is to have the next highest score k. Imagine that an item has four possible scores: 0, 1, 2, and 3. The partial credit model assumes that Equation (6.10) holds between scores 0 and 1, 1 and 2, and 2 and 3:

$$\ln\left[\frac{P(X_{pi}=1)}{P(X_{pi}=0)}\right] = \theta_p - b_{01}\ , \tag{6.11}$$

$$\ln\left[\frac{P(X_{pi}=2)}{P(X_{pi}=1)}\right] = \theta_p - b_{12}\ , \tag{6.12}$$

and

$$\ln\left[\frac{P(X_{pi}=3)}{P(X_{pi}=2)}\right] = \theta_p - b_{23}\ . \tag{6.13}$$

These equations can be combined and generalized for any item i scored $x = 0, ..., m_i$. For the category $x = j$, the partial credit model is

$$P_{ix}(\theta_p) = \frac{\exp\left[\sum\limits_{j=0}^{x}(\theta_p - b_{ij})\right]}{\sum\limits_{r=0}^{m_i}\left[\exp\left[\sum\limits_{j=0}^{r}(\theta_p - b_{ij})\right]\right]} \tag{6.14}$$

as long as

$$\sum_{j=0}^{0}(\theta_p - b_{ij}) \equiv 0 \tag{6.15}$$

(Embretson and Reise 2000; Masters 1982).

In some cases, the Thurstonian thresholds β_j for the different score categories in the partial credit model are used instead of the item step difficulties b_{ij}. The Thurstonian threshold for category j is defined as the ability at which the probability of getting a score less than j equals that probability of getting a score of j or greater (Wilson 2005). The Thurstonian thresholds for our items are used on the Wright Maps shown and discussed in Section 6.4.5 below.

We use the partial credit model for several reasons. First, all of our items are polytomously scored, so the family of dichotomous IRT models are, *prima facie*, not applicable to our data. Second, we deliberately designed our surveys such that each only measures a single construct; thus, we do not need multidimensional IRT models. Finally, the partial credit model allows us to maintain the theoretical advantages the Rasch model has over multi-parameter IRT models. For example, when IRT models add additional item parameters, the relative difficulties of items may change as a function of ability (Wallace and Bailey 2010; Wright 1997). Furthermore, Masters (1988) found that additional parameters may actually mask potential problems with a test's items. For a theoretical comparison of the Rasch and multi-parameter IRT models, see Andrich (2004), Wallace and Bailey (2010), and Wright (1997).

The partial credit model parameters discussed below were estimated using the ConstructMap software. We used the expected *a posteriori* procedure to estimate students' abilities. See Baker and Kim 2004 for a detailed description of IRT parameter estimation techniques.

Section 6.4.2 presents our estimates of the item parameters for Forms A-C. In Sections 6.4.3 and 6.4.4 we examine whether our data meets the assumptions of IRT and whether the partial credit model fits our data, respectively. Section 6.4.5 re-examines the reliabilities of Forms A-C from an IRT perspective.

6.4.2 Items' Difficulties

Table 6.6 shows the step difficulties and Thurstonian thresholds for each item on Forms A-C. Experience suggests that values < -3 logits are abnormally low and values > 3 logits are abnormally high. A cursory glance at Table 6.6 reveals the lowest and highest step difficulties and Thurstonian thresholds for the items of Form A all lie outside of these bounds. The fact that the lowest step difficulties and Thurstonian thresholds are so low is not necessarily a concern. As the scoring rubric in Appendix C shows, the requirements for achieving scores greater than zero are

Table 6.6: The step difficulty b_{ij} and Thurstonian Threshold β_j parameters for the items on Forms A-C for fall 2009. All values are in logits.

		Step Parameters				Thurstonian Thresholds			
	Item	b_{01}	b_{12}	b_{23}	b_{34}	β_1	β_2	β_3	β_4
Form A	Item 1	-6.70	-0.40	2.72	-	-6.76	-0.48	2.71	-
	Item 2	-4.96	0.71	7.45	-	-5.01	0.66	8.14	-
	Item 3	-5.00	1.62	3.68	-	-5.05	1.46	3.75	-
	Item 4	-4.69	1.92	3.49	-	-4.74	1.71	3.60	-
	Item 5	-3.16	-1.52	4.85	-	-3.37	-1.42	4.80	-
Form B	Item 1	-0.68	-0.20	2.77	0.56	-1.05	0.08	1.59	1.88
	Item 2	-0.81	-0.94	2.03	-	-1.34	-0.48	2.08	-
	Item 3	-0.87	-1.35	1.78	-	-1.51	-0.77	1.82	-
	Item 4	-2.11	1.06	1.52	-	-2.16	0.74	1.89	-
	Item 5	-1.44	1.37	-0.97	-	-1.52	0.16	0.42	-
	Item 6	-2.05	1.31	-	-	-2.09	1.34	-	-
Form C	Item 1	-3.62	2.32	-0.92	-	-3.62	0.61	0.80	-
	Item 3	-1.67	1.44	0.50	-	-1.67	1.44	0.49	-
	Item 4	-1.97	1.61	1.56	-	-1.99	1.15	2.06	-
	Item 5	-1.33	2.02	0.06	-	-1.38	0.91	1.25	-

fairly modest. However, the fact that the highest step difficulties and Thurstonian thresholds are all > 3 logits is concerning. This implies that many students are struggling to earn scores of 3 on Form A's items. A score of 3 indicates that students can give the correct answer and provide a complete and correct justification for their answer (see Appendix C). The high values of b_{23} and β_3 are consistent with our CTT analysis that shows many students (including those who used the lecture-tutorials) fail to move into higher score categories post-instruction.

Unlike the step difficulties and Thurstonian thresholds for Form A's items, the parameters in Table 6.6 for Form B and Form C's items do not raise any concerns. However, the reader may wonder how, for example, b_{34} for item 1 on Form B can be larger than b_{23}. Figure 6.3 plots the category response curves as a function of ability for this item. Each curve shows how the probability of earning a particular item score (0, 1, 2, 3, or 4) changes with student ability. As Figure 6.3 shows, the category response curves for scores of 3 and 4 cross at an ability value for which neither are the most probable score. This behavior, which may be somewhat counterintuitive at first glance, is why we use the Thurstonian thresholds instead of the step difficulties on our Wright maps below

(Section 6.4.5).

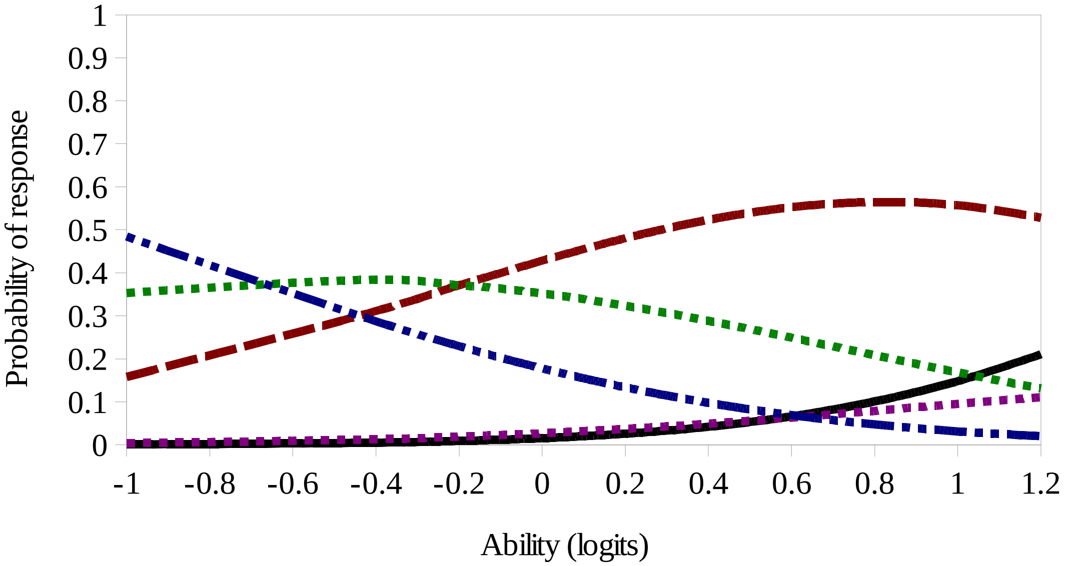


Figure 6.3: The category response curves for item 1, Form B. Each curve shows how the probability of earning a particular score changes with ability. The blue dashed and dotted line corresponds to a score of 0, the green dotted line to a score of 1, the red dashed line to a score of 2, the purple short dashed line to a score of 3, and the solid black line to a score of 4.

6.4.3 Testing Item Response Theory's Assumptions

The ability to test IRT's assumptions and the fit of the model are key advantages IRT possesses over CTT. When the assumptions hold and the model fits the data, one can leverage the strengths of IRT, such as *parameter invariance*. Parameter invariance means that ability estimates do not depend on the specific items administered, and item parameters do not depend on the abilities of respondents (Hambleton and Jones 1993). When the assumptions do not hold and/or when the model does not fit, one cannot claim parameter invariance. Evidence of either is not necessarily fatal to our IRT analysis; since we ultimately only need to make ordinal rankings of the lecture-tutorial and non-lecture-tutorial students in our study, we do not need to leverage all of the advantages IRT possesses over CTT. Furthermore, model misfit and the breakdown of IRT's assumptions *can* signal potential problems in a test and suggest which items are candidates for revision (e.g. Wallace and Bailey 2010). Thus, testing the assumptions and checking model fit are critical aspects of our IRT analysis. In this section we look at IRT's assumptions, while the next section examines model fit.

As noted above, the partial credit model assumes unidimensionality. Tests such as Forms A-C are unlikely to exhibit much of a departure from unidimensionality given their limited number of items. Nevertheless, we must support this assertion with evidence. Smith and Miao (1994) note that item fit statistics, such as the outfit statistic discussed in Section 6.4.4 below, often perform better in detecting departures from unidimensionality in realistic scenarios than more traditional factor analysis methods. Therefore, we use the outfit item fit statistic to detect departures from unidimensionality. Since most of the outfit values for the items on Forms A-C fall within or close to their theoretically expected ranges (see Table 6.10 below), we conclude that we have no evidence for any multidimensionality in Forms A-C.

The partial credit model also assumes local independence. How can one test whether or not this assumption holds? One method uses Yen's Q3 statistic. Yen's Q3 statistic looks at the residuals between the observed and model-predicted responses to each item and then correlates these residuals across respondents by item (Yen 1984). Tables 6.7-6.9 show the Q3 statistic for each pair of items for Forms A-C, respectively.

Yen and Fitzpatrick (2006) recommend flagging all values of the Q3 statistic $\geq |0.20|$. All values exceeding this limit are bolded in Tables 6.7-6.9. Forms A and B each have one item pair with a Q3 value $\geq |0.20|$. On Form A, the Q3 value for the items 3 and 4 is 0.69. This high correlation makes sense given that many students simply reverse their reasoning for item 3 when answering item 4, as discussed in Section 6.5. For Form B, items 1 and 5 have a Q3 statistic of -0.26. We cannot immediately explain why these two items have such a high correlation between their residuals, especially since we do not observe this correlation in data from subsequent semesters (see Chapters 7 and 8). However, the fact that we did not have to flag most of the item pairs on Forms A and B gives us confidence that, on the whole, students' responses to their items exhibit local independence.

We cannot make the same case for Form C. We flagged three item pairs on this four item

Table 6.7: Yen's Q3 statistic for each pair of items on Form A for the fall 2009.

Item	Item 1	Item 2	Item 3	Item 4	Item 5
Item 1	1	0.04	-0.07	-0.11	0.04
Item 2	0.04	1	0.14	0.16	-0.06
Item 3	-0.07	0.14	1	**0.69**	0.03
Item 4	-0.11	0.16	**0.69**	1	0.08
Item 5	0.04	-0.06	0.03	0.08	1

Table 6.8: Yen's Q3 statistic for each pair of items on Form B for the fall 2009.

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Item 1	1	-0.08	0.08	0.18	**-0.26**	0.17
Item 2	-0.08	1	0.19	0.01	0.05	0.07
Item 3	0.08	0.19	1	-0.03	0	0.01
Item 4	0.18	0.01	-0.03	1	0.04	0.13
Item 5	**-0.26**	0.05	0	0.04	1	0.01
Item 6	0.17	0.07	0.01	0.13	0.01	1

Table 6.9: Yen's Q3 statistic for each pair of items on Form C for the fall 2009.

Item	Item 1	Item 3	Item 4	Item 5
Item 1	1	**0.27**	-0.10	**0.20**
Item 3	**0.27**	1	0.14	**0.40**
Item 4	-0.10	0.14	1	-0.03
Item 5	**0.20**	**0.40**	-0.03	1

survey. This indicates that these items require revision, which is exactly what we did for versions of Form C administered during subsequent semesters.

6.4.4 Model Fit

For the Rasch family of IRT models (including the partial credit model), model fit is typically judged by looking at the residuals between the observed and expected scores for a given person and item (Bond and Fox 2001; Wilson 2005; Wu and Adams 2011). Many fit statistics standardize the these residuals by dividing each residual by the standard deviation of student scores on that item (Bond and Fox 2001; Wu and Adams 2011). The outlier sensitive fit (outfit) statistic for items sums the squares of the standardized residuals over the number of respondents n and divides

Table 6.10: Outfit statistics for the items on Forms A-C for the fall 2009.

Item	Form A	Form B	Form C
Item 1	**1.33**	0.99	1.13
Item 2	1.19	0.97	-
Item 3	1.21	0.90	**0.79**
Item 4	1.19	1.16	1.04
Item 5	1.12	1.18	1.02
Item 6	-	1.03	-

this sum by n. Similarly, the outfit statistic for respondents sums the squares of the standardized residuals over the number of items N and divides this sum by N (Bond and Fox 2001; Wu and Adams 2011). In the partial credit model, the outfit statistic for an item indicates how well that item's category response curves match the observed pattern of responses. The outfit statistic for a person indicates how well the responses of that person match the expectations of the model (e.g. whether or not the person correctly answers all items for which she has a high probability of giving the right answer and incorrectly answers all items for which she has a low probability of giving the right answer; Wilson 2005). Both kinds of outfit should have values close to one if the model fits.

How close is close enough? Wu and Adams (2011) note that if, as expected, the squared residuals are χ^2 variates, then item outfits should have variances of $2/n$ and person outfits should have variances of $2/N$. Thus, 95% of items should have outfit values within $1 \pm 1.96\sqrt{2/n}$ and 95% of respondents should have outfit values within $1 \pm 1.96\sqrt{2/n}$ (Wu and Adams 2011).

Table 6.10 shows the outfit statistics for the items on Forms A-C. Using the results of Wu and Adams (2011), we expect items on Form A to have outfit values between 0.81 and 1.19, items on Form B to have outfit values between 0.80 and 1.20, and items on Form C to have outfit values between 0.82 and 1.18. Two items fall outside these values and are highlighted in Table 6.10: Item 1 on Form A and Item 3 on Form C.

Histograms of respondents' outfit values are shown in Figures 6.4-6.6 for Forms A-C, respectively. Following Wu and Adams (2011), we expect 95% of respondents to Form A to have outfit values less than 2.24, 95% of respondents to Form B to have outfit values less than 2.13, and 95%

of respondents to Form C to have outfit values less than 2.39. How many students actually fall within this range? It is 87% for Form A, 93% for Form B, and 100% for Form C. Thus, the model fit appears to be adequate for Forms B and C, but not quite as good as expected for Form A.

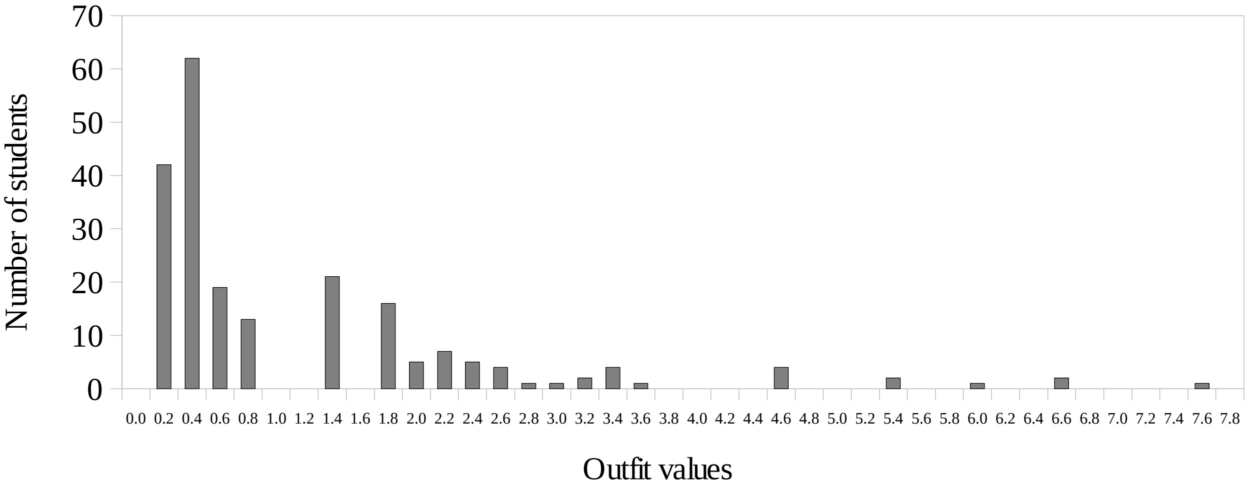


Figure 6.4: A histogram of students' outfit values for Form A for the fall 2009.

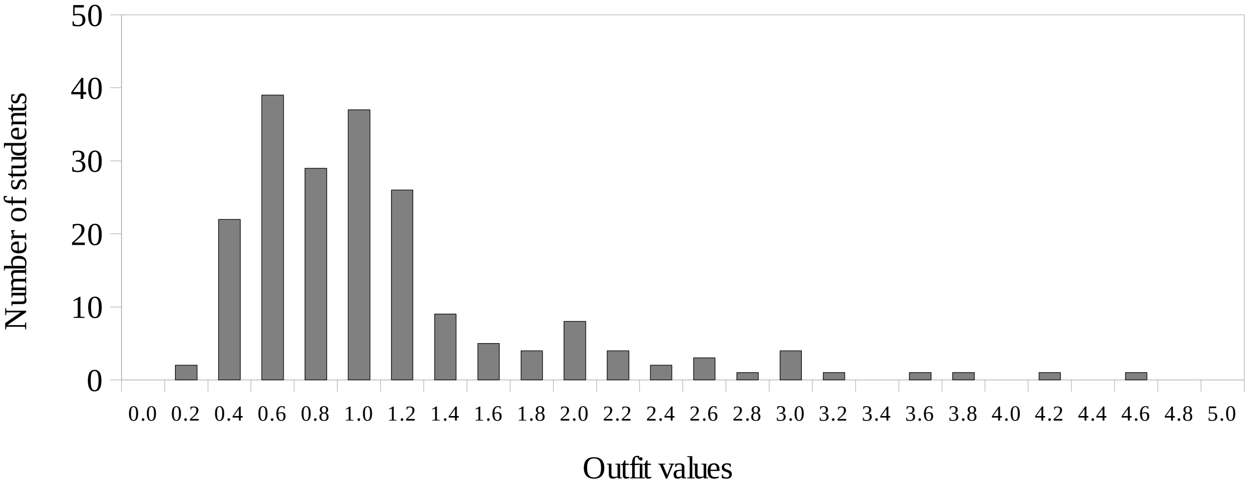


Figure 6.5: A histogram of students' outfit values for Form B for the fall 2009.

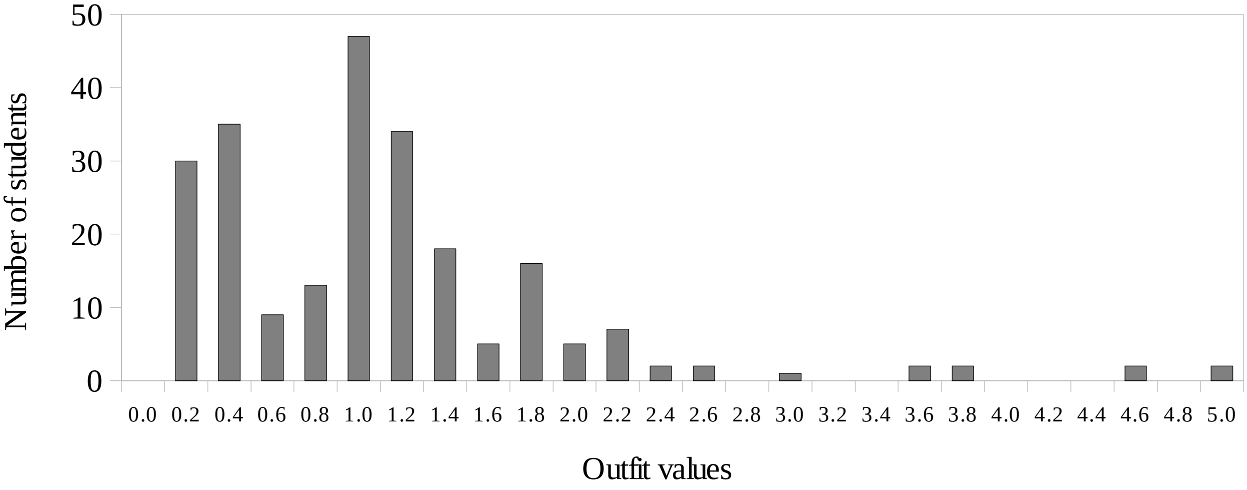


Figure 6.6: A histogram of students' outfit values for Form C for the fall 2009.

6.4.5 Reliability

In section 6.3.3 we examined the reliability of our surveys from a CTT perspective. One estimate of reliability, Cronbach's α, was lower than we hoped for all of our surveys. Unfortunately, the fact that Cronbach's α depends on both the homogeneity of the test taker population and the number of items on each survey (Schmitt 1996; Thompson 2003) left us unable to determine how much these factors contributed to our low values and how much was due to the unreliability of our surveys. The fact that Cronbach's α is always a lower bound on a test's reliability (Lord and Novick 1968) further muddied these waters.

IRT adopts a different approach to addressing the question of reliability. Instead of conceiving of reliability as a single number associated with the entire test, IRT looks at the standard error of measurement of abilities as a function of ability (Embretson and Reise 2000). Furthermore, in the IRT perspective, shorter tests may actually be more reliable than longer tests (Embretson and Reise 2000). The fact that IRT provides a different way to look at test reliability is one of the key reasons we used it to analyze our data.

Figure 6.7 shows how the standard error of measurement of students' abilities changes as a function of ability for Forms A-C. For example, the best estimates of ability Form B provides are around the curve's minimum, or at and around $\theta_p \approx -0.2$. The ability estimates Form B provides become less accurate as one moves away from this value. Figure 6.7 underscores an important point: A reliable survey is one in which the standard error of measurement is smallest in the region where the abilities of most student's lie. Survey designers should thus select items in which the standard error of measurement curve reaches its nadir at the ability at which most students cluster (Hambleton and Jones 1993).

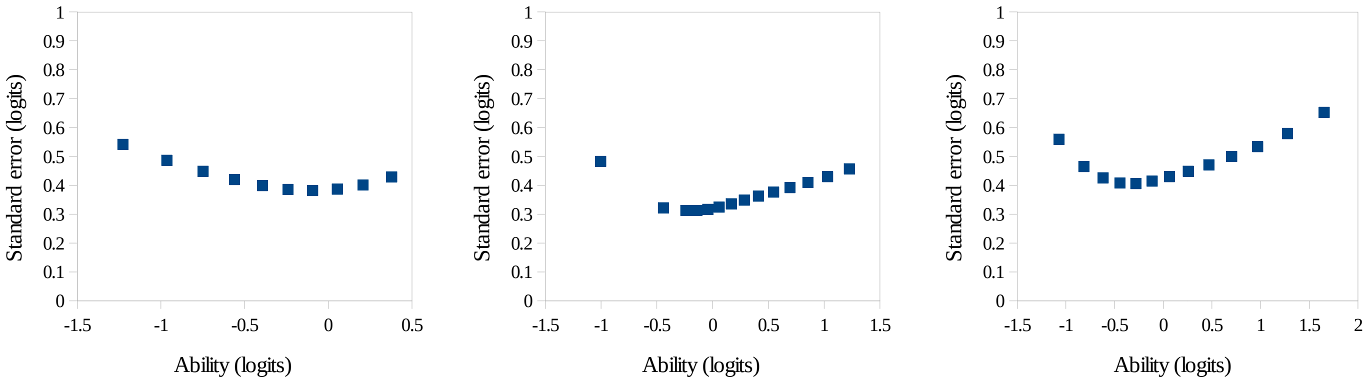


Figure 6.7: Standard error of measurement as a function of ability for (from left to right) Form A, Form B, and Form C for fall 2009.

Do our surveys meet this criterion? One way to tell is to look at the Wright maps (Figures 6.8-6.10) for Forms A-C. Each Wright map has two components. The left part of each Wright map is a histogram of students' estimated abilities (or proficiencies in the nomenclature of ConstructMap, which generated these graphs). The right part shows the logit locations of the Thurstonian thresholds for each item. The logit value at the center of the spread of these Thurstonian thresholds corresponds to the location of the minimum on the associated standard error of measurement plot.

These Wright maps support the hypothesis that Forms A-C are reliable for two reasons. First, the histogram of students' abilities for each form is roughly centered on the ability value at the minimum of the corresponding plot in Figure 6.7. Second, the Thurstonian thresholds "span the space" of students' abilities. This indicates the surveys are reliable since there is not a drastic offset between the logit values covered by the items' Thurstonian thresholds and the location of students' abilities.

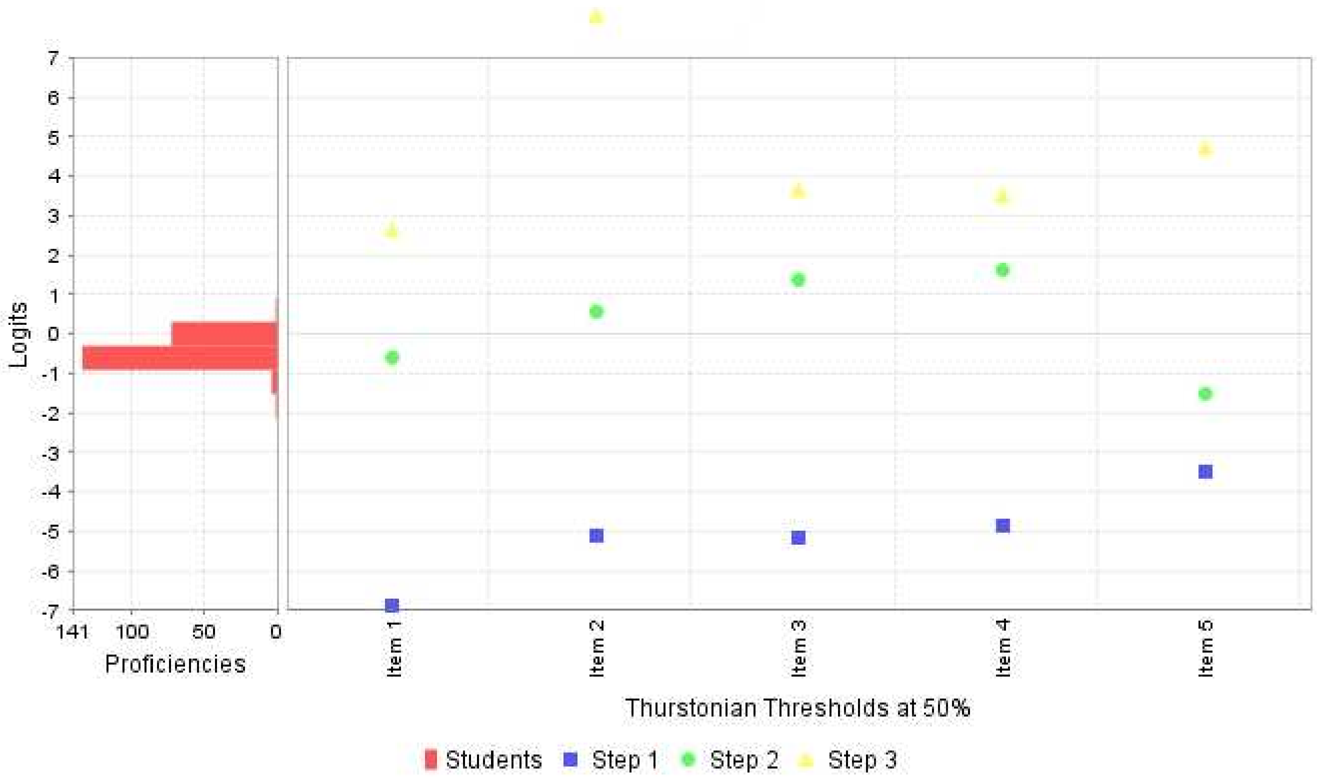


Figure 6.8: The Wright map for Form A for fall 2009.

That being said, the Wright maps also suggest possible improvements for Forms A-C. For example, the offset between students' abilities and the minimum of the standard error of measurement curve is greatest for Form A. This is consistent with the fact that the Thurstonian thresholds for Form A's items are so high that few students are likely to achieve the highest scores. This suggests that the reliability of Form A will improve if we add easier items and/or if we revise Form A's items to make them easier. Similarly, Form C could use items whose highest Thurstonian thresholds align better with the upper tail of the distribution of students' abilities. These issues are reconsidered in Section 6.7.

The Wright maps can also, in principle, help us connect this measurement model to our construct maps in Chapter 4 (Wilson 2005). In those construct maps, we postulated three levels on the *Hubble plots* construct (Form A), four levels on the *models* construct, and three levels on the *evolving universe* construct. One way to connect these construct maps to the Wright maps is to

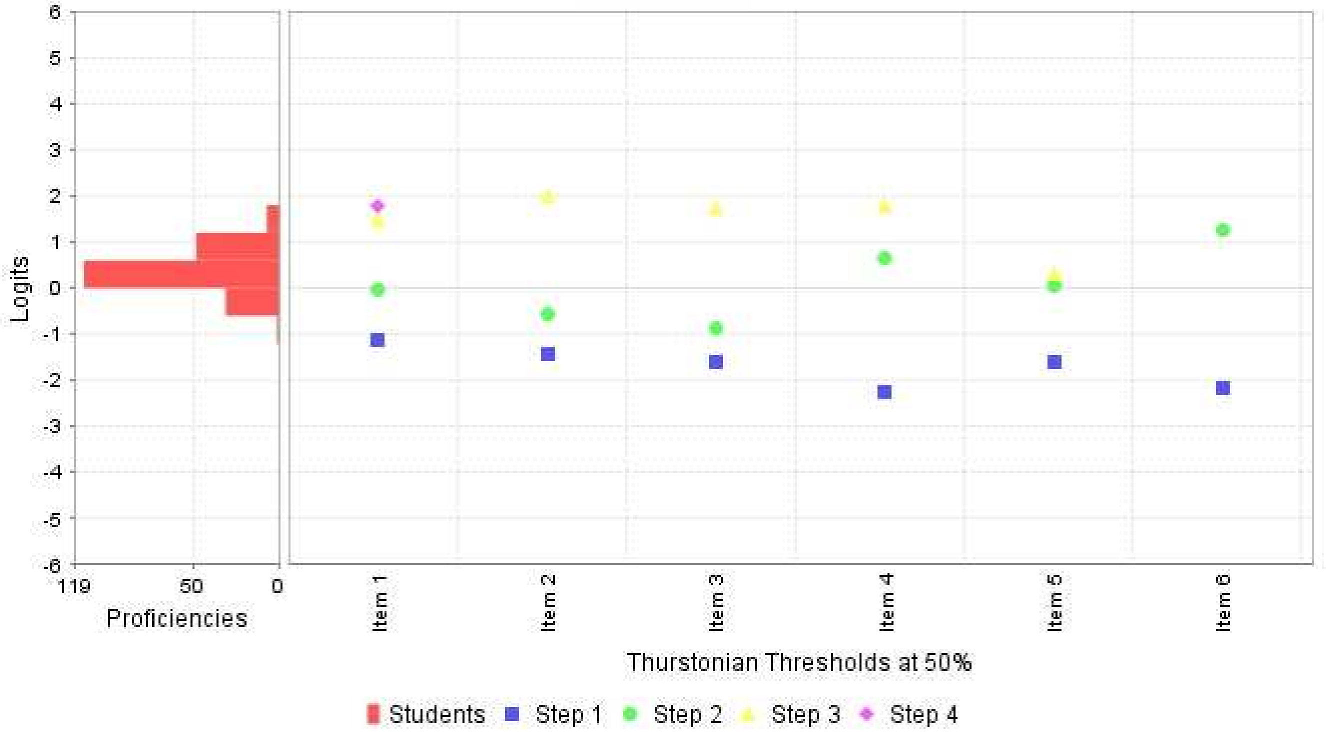


Figure 6.9: The Wright map for Form B for fall 2009.

define "cut points" in ability that separate one construct map level from another. For example, we could use Form C's Wright map (Figure 6.10) to define Level 1 on the *evolving universe* construct map as all abilities below -1 logit, Level 2 as all abilities between -1 and 0.5 logits, and Level 3 as all abilities above 0.5 logits. However, this procedure is somewhat arbitrary. Furthermore, the lowest standard error of measurement for any ability on Form C is 0.41 logits, which means our estimates of abilities using Form C are, at best, accurate within ±0.80 logits with 95% confidence. This calls into question our ability to accurately place students at a given level on the *evolving universe* construct map. Similar concerns apply to the other Wright maps and construct maps.

Of course, we are not ultimately concerned with being able to place individual students at different levels on the construct maps. We care about whether or not students who use the lecture-tutorials display, on average, higher learning gains than students who do not. Our partial credit model estimates of students' abilities are adequate for this task, as described in the following

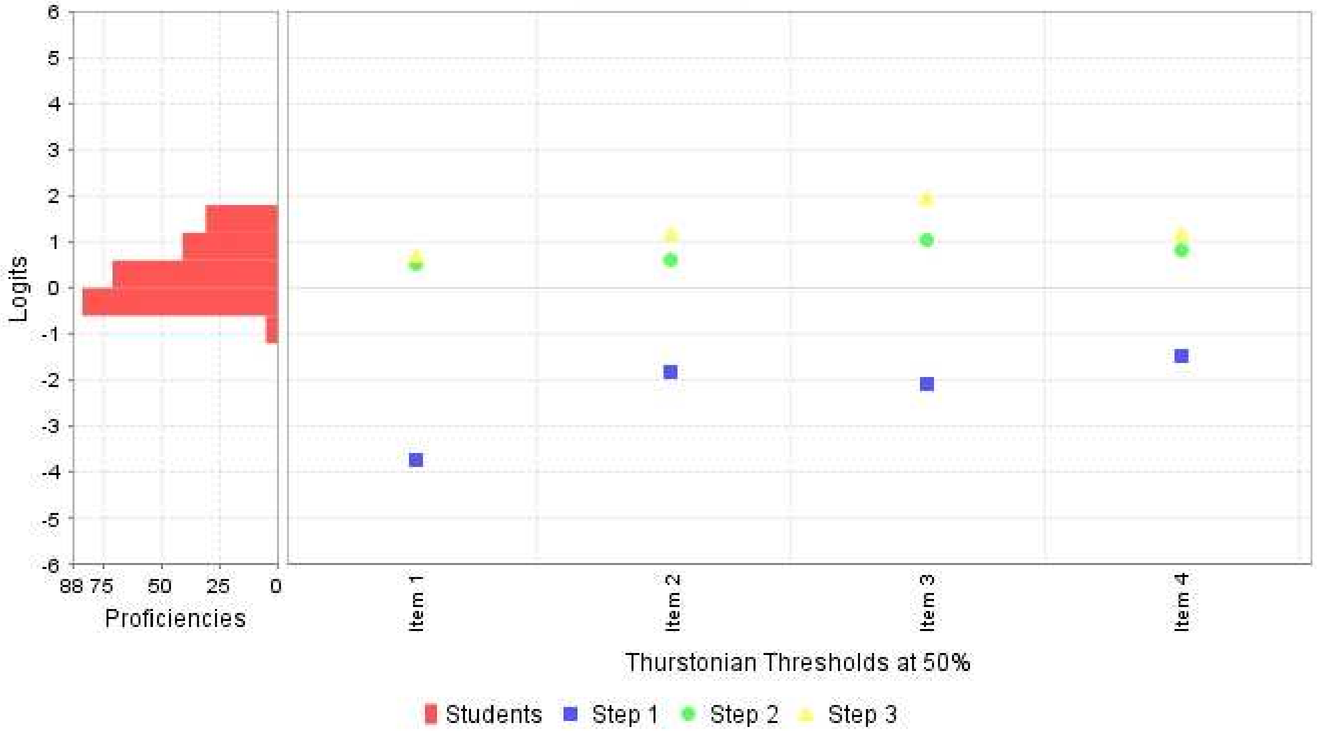


Figure 6.10: The Wright map for Form C for fall 2009.

subsection.

6.4.6 IRT Gains

In Section 6.3.4 we calculated the normalized gains for the lecture-tutorial and non-lecture-tutorial populations of students. We also used the Mann-Whitney test to compare the pre- and post-instruction scores of the lecture-tutorial students and the pre- and post-instruction scores of the non-lecture-tutorial students. This test allowed us to check whether or not the normalized gains are statistically significant. In this section, we compare the gains of the lecture-tutorial and non-lecture-tutorial students on Forms A-C using their IRT estimates of abilities.

The first step in this comparison is calculating the average abilities of both lecture-tutorial and non-lecture-tutorial students, pre- and post-instruction. For n students, the weighted average

Table 6.11: Average pre-instruction IRT scores, post-instruction IRT scores, and IRT gains for the lecture-tutorial and non-lecture-tutorial classes, as well as the difference between their gains, for Forms A-C in the fall 2009. All values are in logits.

Form	LT pre	Non-LT pre	LT post	Non-LT post	LT Gain	Non-LT Gain	LT Gain-Non-LT Gain
Form A	-0.36 ± 0.04	-0.57 ± 0.12	-0.31 ± 0.04	-0.32 ± 0.11	0.05 ± 0.06	0.25 ± 0.16	-0.20 ± 0.17
Form B	0.10 ± 0.04	0.29 ± 0.07	0.58 ± 0.04	0.43 ± 0.10	0.48 ± 0.06	0.15 ± 0.12	0.34 ± 0.14
Form C	-0.09 ± 0.04	0.15 ± 0.08	0.70 ± 0.06	0.23 ± 0.09	0.79 ± 0.07	0.09 ± 0.12	0.70 ± 0.14

ability is

$$\hat{\theta} = \frac{\sum\limits_{p=1}^{n} w_p \theta_p}{\sum\limits_{p=1}^{n} w_p} \,, \tag{6.16}$$

where the weights w_p are related to the abilties' standard errors of measurement σ_p by

$$w_p = 1/\sigma_p \tag{6.17}$$

(Taylor 1982). The uncertainty in the weighted average is

$$\hat{\sigma} = \left(\sum_{p=1}^{n} w_p \right)^{-1/2} \tag{6.18}$$

(Taylor 1982). The first four columns of Table 6.11 show the average abilities and their uncertainties for both the lecture-tutorial and non-lecture-tutorial students, pre- and post-instruction.

Following Wallace and Bailey (2010), we calculated the gains for each group of students as simply the average post-instruction ability minus the average pre-instruction ability. We found the uncertainties in the gains by adding the uncertainties in the pre- and post-instruction scores in quadrature. These gains and their uncertainties are also shown in Table 6.11.

Finally, we subtracted the non-lecture-tutorial gains from the lecture-tutorial gains. Once again, we computed the uncertainty in this difference by adding the uncertainties of the gains in quadrature. The differences in gains are given in the final column of Table 6.11.

As Table 6.11 demonstrates, we cannot claim that the lecture-tutorials helped students improve their performances on Form A. However, the gains of the lecture-tutorial students are significantly larger than the gains of the non-lecture-tutorial students on Forms B and C. On Form B, the difference in gain is 2.4 times larger than its uncertainty, while on Form C the difference is 5 times larger than its uncertainty. These results support the idea that the lecture-tutorials improve students' performances on the constructs measured by Forms B and C.

6.5 Breakdown of Item Responses

In Sections 6.3 and 6.4, we examined our results from the perspectives of classical test theory and item response theory, respectively. These results provide evidence that the lecture-tutorials are affecting measurable improvements on students' performances on Forms B-D, but not on Form A. However, we have yet to explore what students are actually saying in their responses to the items on Forms A-D. This section fills in this gap.

In many of the tables and figures below, we present the results for individual classes in addition to aggregate results. I will not comment too much on the results of and differences between individual classes, since many of these numbers are likely tied to demographic and implementation factors that lie beyond the scope of this study. They are primarily included in order to provide interested readers an overview of how each class performed. Our analysis focuses on overall patterns in students responses.

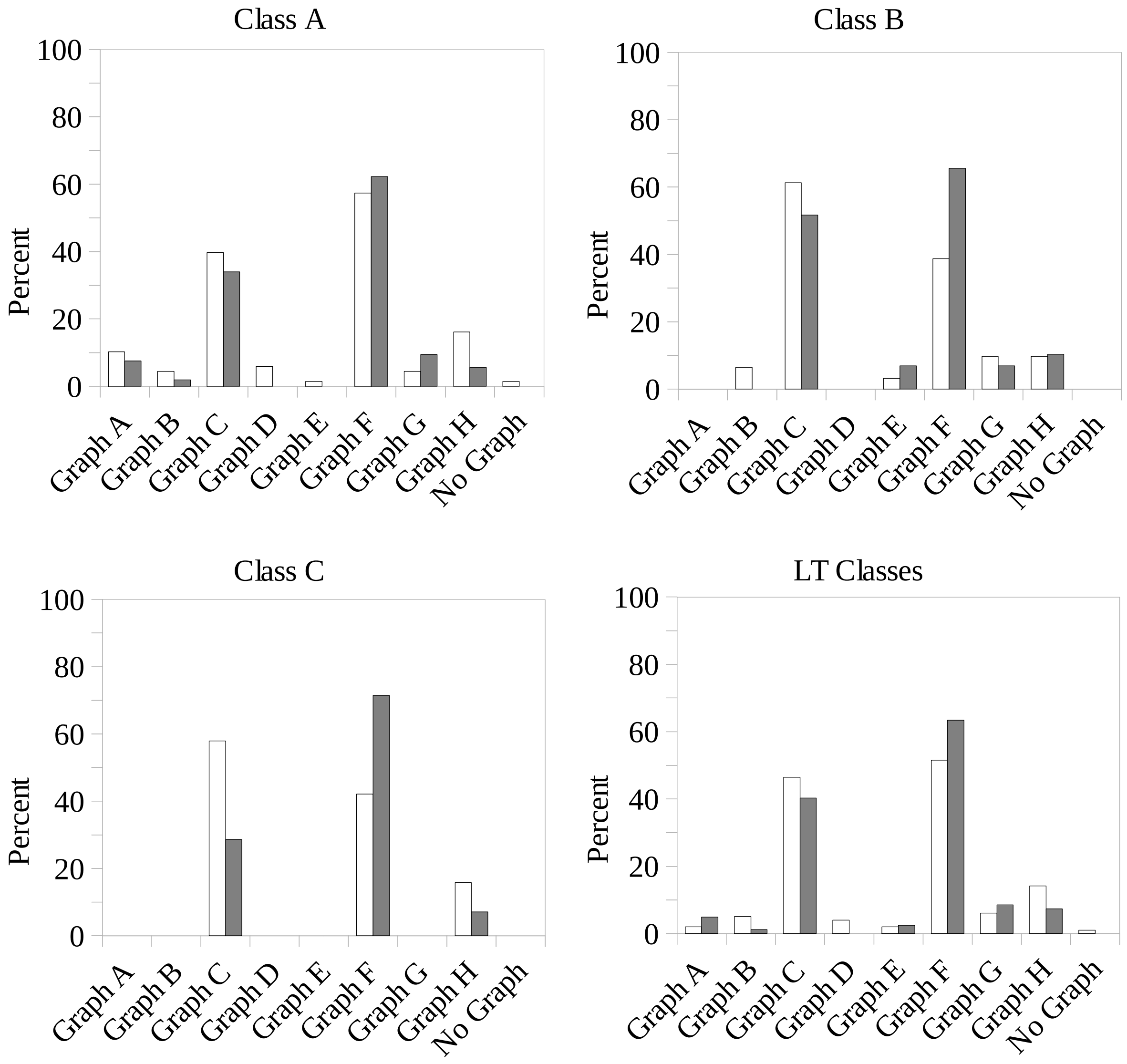
6.5.1 Form A Responses

Each student's response to each item on Form A received an overall score between 0 and 3 (see Appendix C). Table 6.12 shows the percent of students with overall scores of 0, 1, 2, and 3 for Classes A-C as well as for the lecture-tutorial classes (Classes A and B) combined. Table 6.12 shows that few students earned scores of 0 or 3 for any item, pre- and post-instruction. This result holds across all three classes. Pre- and post-instruction, the vast majority of students earned scores of 1 or 2 which, according to the scoring rubric in Appendix C, means that most students either gave incorrect answers or correct answers that were chosen for incorrect or incomplete reasons. The fact that Table 6.12 shows little improvement for both the lecture-tutorial classes as well as Class C is consistent with the results cited in Sections 6.3.2, 6.3.4, 6.4.5, and 6.4.6 and implies that the lecture-tutorials did not measurably improve students' performances on Form A.

Table 6.12: Pre- and Postinstruction distribution of scores on Form A for fall 2009.

		Class A				Class B				Class C				LT Classes			
Preinstruction	Item	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
	Item 1	1%	60%	35%	3%	0%	68%	29%	6%	0%	74%	21%	5%	1%	65%	32%	2%
	Item 2	4%	69%	26%	0%	0%	87%	13%	0%	5%	79%	16%	0%	3%	77%	20%	0%
	Item 3	4%	88%	7%	0%	3%	94%	3%	0%	0%	100%	0%	0%	4%	90%	6%	0%
	Item 4	3%	90%	7%	0%	10%	87%	3%	0%	0%	100%	0%	0%	5%	89%	6%	0%
	Item 5	12%	25%	63%	0%	13%	10%	71%	3%	5%	16%	79%	0%	12%	20%	67%	1%
Postinstruction		Class A				Class B				Class C				LT Classes			
	Item	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
	Item 1	0%	49%	49%	2%	0%	54%	48%	0%	0%	36%	57%	7%	0%	51%	46%	2%
	Item 2	0%	77%	25%	0%	7%	75%	19%	0%	0%	71%	29%	0%	2%	76%	22%	0%
	Item 3	0%	66%	32%	2%	7%	93%	0%	0%	0%	100%	0%	0%	2%	76%	21%	1%
	Item 4	4%	72%	23%	2%	3%	96%	0%	0%	0%	100%	0%	0%	4%	80%	15%	1%
	Item 5	2%	45%	55%	0%	10%	54%	37%	0%	7%	43%	50%	0%	5%	49%	46%	0%

Figures 6.11-6.14 show the percentage of students in each class (as well as for the lecture-tutorial classes combined) that selected each graph choice for items 1-4, respectively. Note that the percentages may not add up to 100% since some students selected multiple graphs for a single item.


Class A

Class B

Class C

LT Classes

Figure 6.11: Students' graph choices for item 1 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

For item 1, graph F is the correct answer. As Figure 6.11 shows, graphs C and F are the most popular choices. Why are students drawn to these two choices? The overwhelming majority of students who selected graph C did so because the line is horizontal: Among the lecture-tutorial students, 72% (pre-instruction) and 57% (post-instruction) justified their choice by the fact that the line and/or the velocity is constant. For the non-lecture-tutorial students, these numbers are 82% (pre-instruction) and 100% (post-instruction). These results suggest that students who select graph C do so because they equate “expansion at a constant rate” with “constant velocity.”

What about the students who selected graph F? Many discussed relevant features of the graph in their responses. For example, pre-instruction (post-instruction), 24% (23%) of the lecture-tutorial students and 13% (10%) of the non-lecture-tutorial students said graph F shows the velocity increasing. Pre-instruction (post-instruction), 35% (29%) of the lecture-tutorial students and 13% (10%) of the non-lecture-tutorial students talked about how graph F shows the distance increasing. Finally, a number of students (12% pre-instruction and 21% post-instruction for the lecture-tutorial group and 25% pre-instruction and 40% post-instruction for the non-lecture-tutorial group) discussed the constant slope of graph F in their responses. Despite the fact that many students highlighted relevant features of graph F in their answers, few put these features together to form a complete and correct response and earn an overall score of 3.

For item 2, the most frequently selected graphs are B (the correct answer), E, and G. As was the case for students who chose graph F for item 1, students who chose graph B for item 2 often highlighted relevant features of the graph but failed to combine their knowledge of these features into a complete and correct justification. Furthermore, students who selected graphs B and E tended to focus on the fact that some aspect of the graph was decreasing or negative. For example, 64% of the lecture-tutorial students pre-instruction chose graph B because “it” or “the velocity” is decreasing. After instruction, this changed to 50%. This was the reasoning of 31% (40%) of the non-lecture-tutorial students pre-instruction (post-instruction). For the lecture-tutorial students, 30% (26%) of students who chose graph E used similar reasoning pre-instruction (post-instruction), while for the non-lecture-tutorial students it was 20% (25%). What about the students who said the answer to item 2 is graph G? Like the students who chose graph C for item 1, these students typically focused on the velocity: Of the lecture-tutorial students who selected graph G, 48% (30%) said pre-instruction (post-instruction) that the velocity is constant, while 35% (26%) said the velocity is negative. For the non-lecture-tutorial students, 50% (100%) pre-instruction (post-instruction) talked about graph G's constant velocity and 25% (67%) talked about how it is negative. Overall, the reasoning patterns employed by students answering item 2 are similar to the reasoning patterns used by students answering item 1.

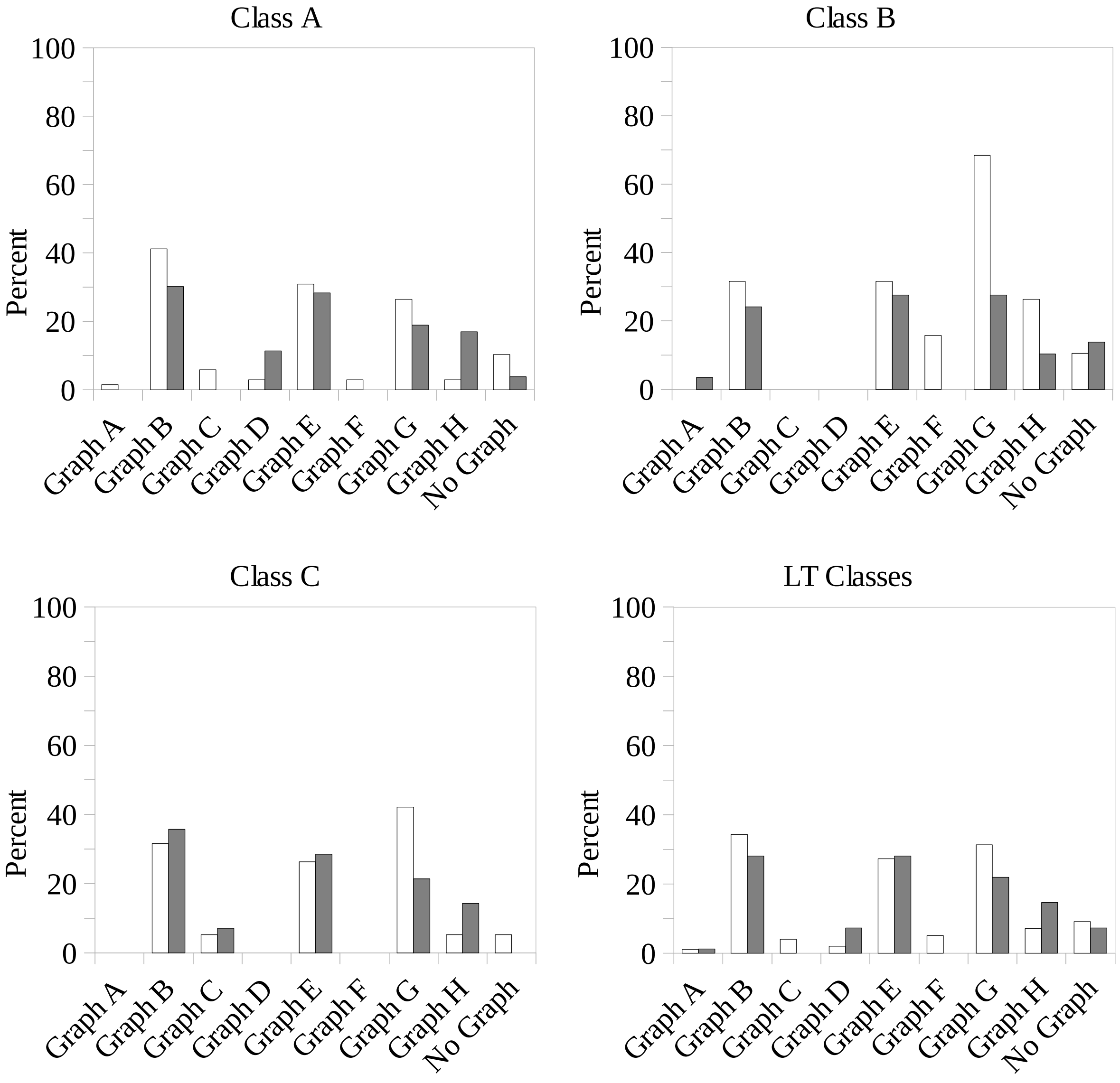


Figure 6.12: Students' graph choices for item 2 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Most students either choose graph D or graph F as their answer for item 3, even though

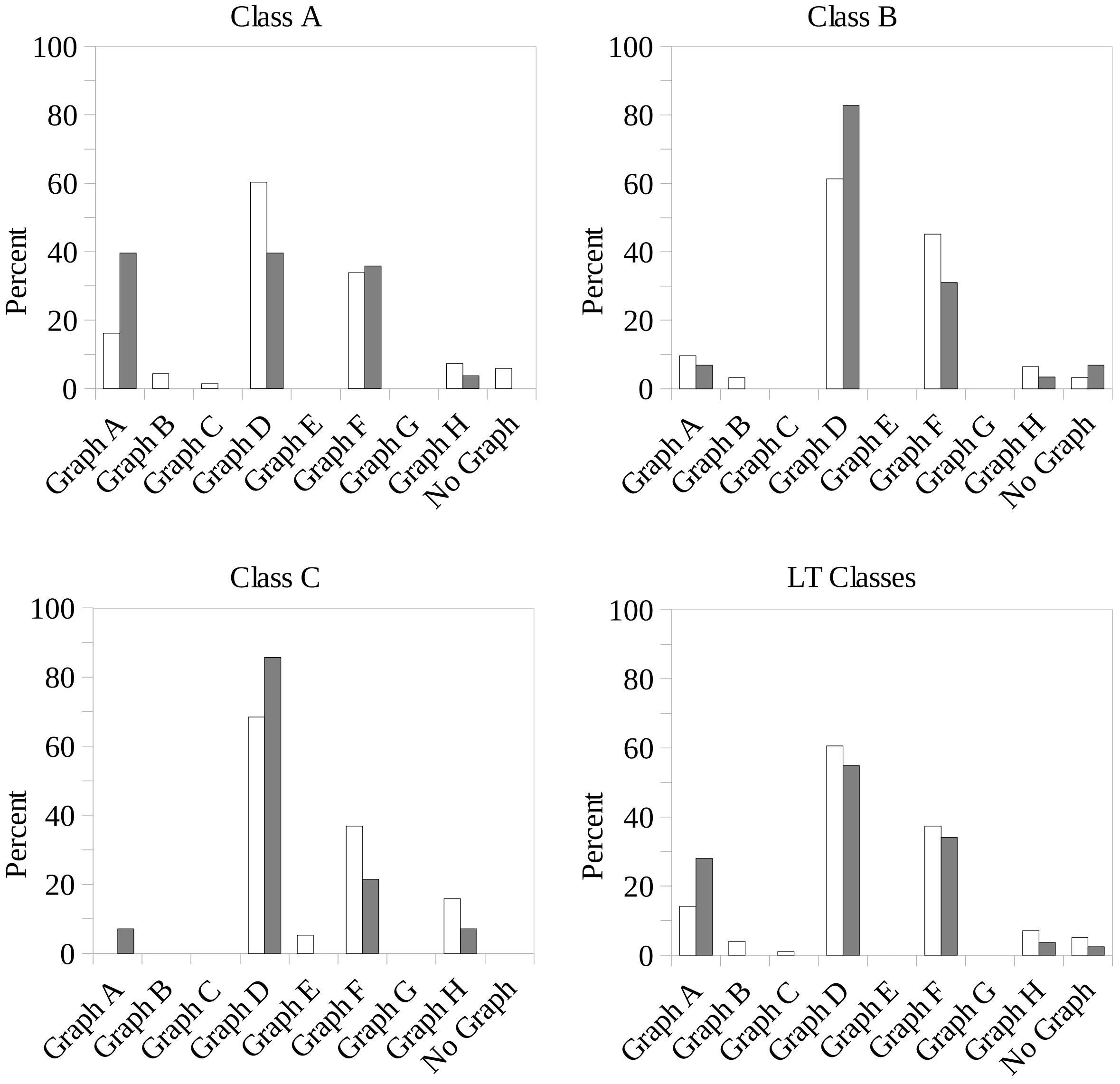


Figure 6.13: Students' graph choices for item 3 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

graph A is the correct answer. The responses of many students indicate that they are making one of the common graph interpretation errors noted by McDermott, Rosenquist, and van Zee (1987): They answer this item by referring to the height of the line rather than its slope. For example, 68% (33%) of lecture-tutorial students pre-instruction (post-instruction) say graph D shows a universe expanding at a faster and faster rate over time because it shows the velocity or the line increasing. For the non-lecture-tutorial students, these percentages are 69% pre-instruction and

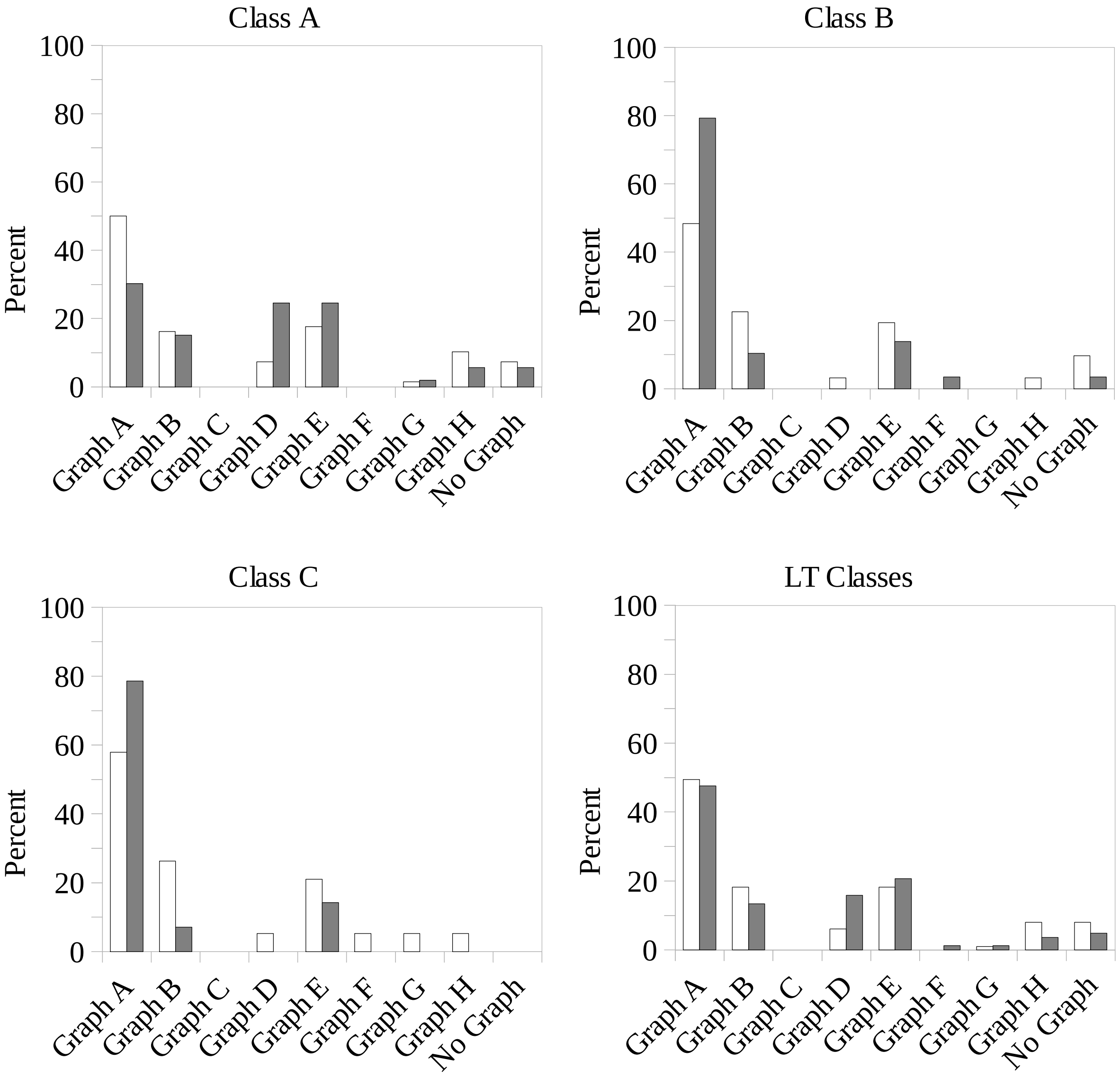


Figure 6.14: Students' graph choices for item 4 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

50% post-instruction. Similarly, the 68% (71%) of the lecture-tutorial students who selected graph F pre-instruction did so because it shows the velocity increasing. 71% (67%) of non-lecture-tutorial students used the same reasoning pre-instruction (post-instruction). Similarly, most students say, in item 4, that graph A represents a universe expanding at a slower and slower rate over time based on the fact that the line levels off at increasing distance.

What about item 5? Figure 6.15 shows the distribution of students' graph selections pre-and post-instruction. Note that although graph C is the correct answer, the percentage of students selecting graph C actually declined pre- to post-instruction. As I discuss in Section 6.6, student interviews indicate that many students misinterpreted what this item is asking. This is consistent with our analysis of students' written responses. Because many students answered a different question than we intended for item 5, we will not discuss it any further in this section since it did not provide much useful information on students difficulties interpreting Hubble plots.

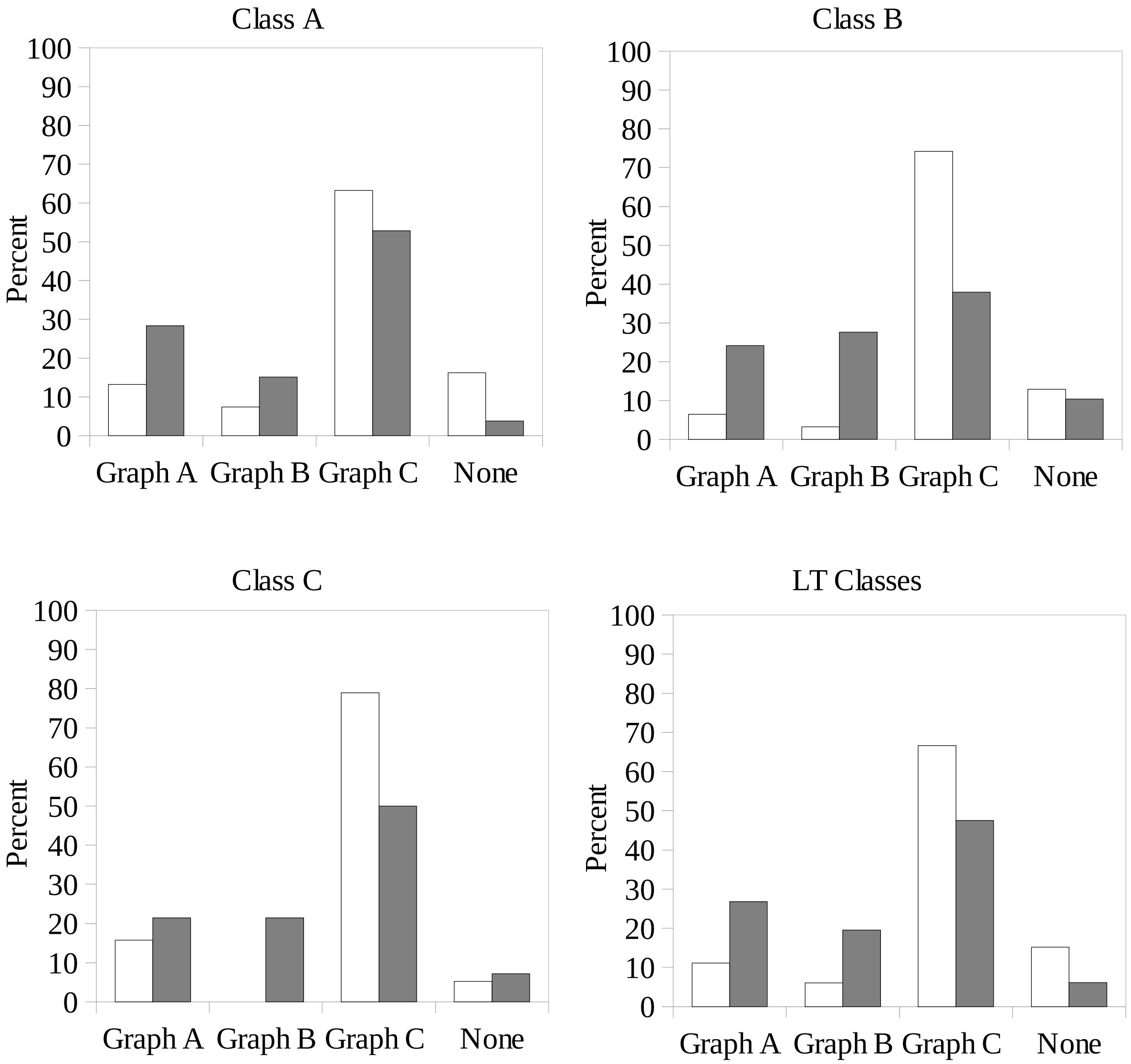


Figure 6.15: Students' graph choices for item 5 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

6.5.2 Form B Responses

Table 6.13 shows the percentage of students with each of the possible overall item scores for each class as well as for the lecture-tutorial classes combined. Unlike Form A (Table 6.12), a comparison of the pre- and post-instruction scores shows a noticable shift on many items from lower to higher scores. This effect appears to be stronger for the two lecture-tutorial classes, Classes A and B, than for Class C, which did not use the lecture-tutorials.

What, specifically, are students writing when they respond to Form B's items? Table 6.14 shows the percent of students who use each reasoning element for item 1 listed in the scoring rubric (Appendix C) for each class, pre- and post-instruction. The percents may not sum to 100% since students often used multiple reasoning elements in constructing their responses.

The data in Table 6.14 shows an interesting result. A number of students pre-instruction claim that the "expansion of the universe" is a metaphor for how our knowledge of the universe increases over time and/or for how new objects are created over time. For example, one student wrote the following:

> "I dont [*sic*] think that it is acculuy [*sic*] expanding in a physical sense, but instead our knowledge of the univers [*sic*] and the areas that we have discovered is expanding with an increase in technology and investments in sciens [*sic*]."

These ideas about expansion are not prevalent in the post-instruction results for both the lecture-tutorial and non-lecture-tutorial students. This appears to be a case where traditional instruction is as effective as the lecture-tutorials in moving students away from this incorrect idea. In fact, Table 6.13 shows that Class C was about as effective as the lecture-tutorial classes in improving students overall scores on item 1.

We cannot tell the same story for item 2. The lecture-tutorial students show a very different pattern of responses than the non-lecture-tutorial students. As shown in Table 6.15, pre-instruction, only about 10% of lecture-tutorial and non-lecture tutorial students connect the Big Bang to the beginning of the expansion of the universe. Post-instruction, this improves to 29% for the non-lecture-tutorial students and 50% for the lecture-tutorial students. Pre-instruction, over half of

Table 6.13: Pre - and Post-instruction distribution of scores on Form B for fall 2009.

		Class A					Class B					Class C					LT Classes				
Preinstruction	Item	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
	Item 1	9%	51%	37%	1%	1%	15%	10%	70%	5%	0%	5%	14%	59%	9%	14%	10%	42%	44%	2%	1%
	Item 2	9%	34%	57%	0%	-	5%	10%	80%	5%	-	0%	0%	95%	4%	-	8%	29%	62%	1%	-
	Item 3	3%	24%	61%	11%	-	10%	5%	70%	15%	-	0%	5%	68%	27%	-	4%	20%	63%	12%	-
	Item 4	4%	76%	11%	9%	-	15%	60%	10%	15%	-	5%	45%	27%	23%	-	7%	72%	11%	10%	-
	Item 5	3%	30%	9%	59%	-	5%	50%	5%	40%	-	5%	41%	27%	27%	-	3%	34%	8%	54%	-
	Item 6	4%	69%	27%	-	-	10%	65%	25%	-	-	9%	50%	41%	-	-	6%	68%	28%	-	-
		Class A					Class B					Class C					LT Classes				
Postinstruction	Item	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
	Item 1	5%	2%	46%	14%	32%	22%	6%	39%	6%	28%	0%	0%	57%	14%	29%	9%	3%	45%	12%	31%
	Item 2	2%	2%	41%	55%	-	6%	0%	61%	33%	-	7%	0%	71%	21%	-	3%	1%	46%	50%	-
	Item 3	4%	0%	54%	43%	-	0%	0%	61%	39%	-	7%	0%	64%	29%	-	3%	0%	55%	42%	-
	Item 4	0%	18%	52%	30%	-	0%	11%	50%	39%	-	7%	21%	57%	14%	-	0%	16%	51%	32%	-
	Item 5	0%	9%	7%	84%	-	17%	22%	6%	56%	-	21%	0%	14%	64%	-	4%	12%	7%	77%	-
	Item 6	2%	50%	48%	-	-	6%	67%	28%	-	-	7%	36%	57%	-	-	3%	54%	43%	-	-

Table 6.14: Common reasoning elements used by students in their answers to Form B, Item 1: Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding."

	Preinstruction				Postinstruction			
Reasoning Element	Class A	Class B	Class C	LT Classes	Class A	Class B	Class C	LT Classes
Student says the size of the universe increases over time.	17%	20%	18%	18%	13%	22%	7%	15%
Student says the universe has a center.	7%	15%	14%	9%	0%	6%	0%	1%
Student says the universe has no center.	0%	0%	0%	0%	5%	0%	0%	4%
Student says the universe has an edge.	1%	5%	0%	2%	0%	6%	0%	1%
Student says the universe has no edge.	4%	5%	5%	4%	4%	0%	0%	3%
Student talks about redshifts/Doppler shifts.	1%	5%	14%	2%	16%	0%	0%	12%
Student says space(time) is growing/stretching.	3%	0%	5%	2%	64%	17%	29%	53%
Student talks about the movement of galaxies and/or their increasing distances.	3%	15%	27%	6%	34%	44%	36%	36%
Student talks about the movement of stars and/or their increasing distances.	9%	15%	9%	10%	0%	6%	0%	1%
Student talks about the movement of planets and/or their increasing distances.	7%	10%	9%	8%	0%	0%	0%	0%
Student talks about the movement of "objects" (something unspecified or not a star, planet, or galaxy) and/or their increasing distances.	7%	15%	32%	9%	23%	6%	29%	19%
Student says the distances between everything increase.	6%	30%	14%	11%	14%	11%	7%	14%
Student says farther objects move away faster.	1%	0%	9%	1%	5%	0%	14%	4%
Student talks about the Big Bang.	9%	40%	27%	16%	2%	22%	14%	7%
Student talks about an explosion.	4%	10%	14%	6%	0%	6%	0%	1%
Student says the early universe was once hot, small, and/or dense.	1%	0%	0%	1%	14%	6%	0%	12%
Student says we learn more about the universe over time.	30%	5%	5%	24%	2%	0%	7%	1%
Student talks about how we are looking further back in time as we look farther in space.	4%	0%	9%	3%	0%	6%	7%	1%
Student says new things are created in the universe over time.	23%	5%	5%	19%	0%	6%	0%	1%
Student gives irrelevant over time.	0%	10%	5%	2%	4%	17%	0%	7%
Student gives some other reason not specified above.	10%	5%	0%	9%	4%	11%	21%	5%
Student says s/he has no idea.	6%	0%	0%	4%	0%	0%	0%	0%
Answer field is blank or the student provided no reason.	1%	5%	0%	2%	0%	0%	7%	0%

Table 6.15: Common reasoning elements used by students in their answers to Form B, Item 2: Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory."

	Preinstruction				Postinstruction			
Reasoning Element	Class A	Class B	Class C	LT Classes	Class A	Class B	Class C	LT Classes
Student says the Big Bang is the beginning of the universe.	29%	55%	32%	34%	32%	33%	21%	32%
Student says the Big Bang is the beginning of expansion.	13%	10%	9%	12%	50%	50%	29%	50%
Student says the Big Bang was the beginning of something smaller than the universe.	23%	10%	0%	20%	4%	0%	0%	3%
Student says the Big Bang is an event that happened to something smaller than the universe.	13%	0%	0%	10%	0%	6%	0%	1%
Student says the Big Bang is the beginning of space.	1%	0%	0%	1%	4%	6%	0%	4%
Student says the Big Bang is the beginning of time.	3%	10%	0%	4%	14%	11%	7%	14%
Student talks about the creation/production of elements.	0%	15%	32%	3%	0%	17%	21%	4%
Student says the Big Bang was an explosion.	46%	70%	59%	51%	17%	39%	50%	23%
Student says the Big Bang was not an explosion.	0%	0%	0%	0%	2%	0%	0%	1%
Student says matter existed before the Big Bang.	27%	35%	45%	29%	14%	17%	43%	15%
Student says there was a dense piece of matter before the Big Bang.	7%	0%	18%	6%	7%	11%	29%	8%
Student talks about matter coming together before the Big Bang.	10%	10%	18%	10%	0%	0%	14%	0%
Student says matter formed from energy.	0%	0%	9%	0%	32%	11%	0%	27%
Student says the early universe was hot, dense, and/or small.	1%	0%	9%	1%	63%	39%	29%	57%
Student says the Big Bang was an event that happened in empty space.	0%	0%	0%	0%	0%	0%	0%	0%
Student gives irrelevant information.	4%	0%	0%	3%	2%	0%	0%	1%
Student gives some other reason not specified above.	9%	5%	5%	8%	5%	17%	0%	8%
Student says s/he has no idea.	6%	0%	0%	4%	2%	0%	0%	1%
Answer field is blank or the student provided no reason.	0%	5%	0%	1%	0%	0%	7%	0%

the student responses talk about the Big Bang as an explosion. After instruction, 50% of the non-lecture-tutorial students still describe the Big Bang as an explosion, but only 23% of the lecture-tutorial students do the same. Finally, a significant minority of students (45% of non-lecture-tutorial students and 29% of lecture-tutorial students) talk about matter existing before the Big Bang pre-instruction. After instruction, these percentages change to 43% and 15%, respectively. These results are consistent with previous papers that claim that students conceive of the Big Bang as the explosion of pre-existing matter into empty space (Prather, Slater, and Offerdahl 2003; Lineweaver and Davis 2005), and they support the hypothesis that the cosmology lecture-tutorials help students develop more expert-like ideas about the Big Bang.

Note that Table 6.15 also shows that some students think the Big Bang refers to the beginning of something smaller than the universe or an event that occurred to something smaller than the universe. Responses stating that the Big Bang was the start of the galaxy, Solar System, or Earth fall into the former category. The most common response for the latter category is that the Big Bang refers to the asteroid that struck the Earth at the end of the Mesozoic Era 65 million years ago. Both categories of responses are more common pre-instruction than post-instruction and both are consistent with previous studies (Prather, Slater, and Offerdahl 2003; Simonelli and Pilachowski 2004).

Table 6.16 gives the percent of students who used each of the listed reasoning elements in their responses to item 3. The responses to this item are hard to interpret. Many reflect the fact that many students may not have interpreted the item as we intended. Ideally, we wanted students to answer "no" to this item because there would have been no location in the universe outside of the Big Bang. We expected students to answer "yes" if they thought the Big Bang happened at a specific location in the universe and was initially surrounded by empty space. Responses such as "the Big Bang was not safe" or "the Big Bang would not have been visible to the naked eye" do not help us determine whether or not a student understands that the entire universe participated in the Big Bang. The same difficulty plagues our ability to make sense of responses in the category "there was nothing/no locations before the Big Bang." For example, consider the following student

Table 6.16: Common reasoning elements used by students in their answers to Form B, Item 3: Imagine you had a spaceship that could take you back in time to the time of the Big Bang. Would there have been any locations in the universe from which you could have watched the Big Bang from a distance?

	Preinstruction				Postinstruction			
Reasoning Element	Class A	Class B	Class C	LT Classes	Class A	Class B	Class C	LT Classes
Student talks about how the universe used to be small and/or how there was no space in the universe outside the Big Bang.	14%	15%	27%	14%	45%	61%	50%	49%
Student says there was space in the universe outside the Big Bang.	17%	25%	9%	19%	7%	11%	0%	8%
Student says there was no time before the Big Bang.	1%	5%	0%	2%	0%	6%	0%	1%
Student says there was time before the Big Bang.	0%	0%	0%	0%	0%	0%	7%	0%
Student says there was no matter before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was matter before the Big Bang.	3%	5%	9%	3%	0%	17%	29%	4%
Student says the universe began at the Big Bang.	13%	15%	18%	13%	9%	6%	7%	8%
Student says there was nothing/no locations before the Big Bang.	21%	25%	18%	22%	21%	22%	14%	22%
Student says the Big Bang was an explosion.	6%	5%	14%	6%	4%	6%	0%	4%
Student says the Big Bang was not safe.	10%	10%	14%	10%	2%	6%	0%	3%
Student talks about high temperatures.	5%	0%	0%	3%	13%	17%	14%	14%
Student talks about high density.	0%	0%	0%	0%	11%	6%	0%	9%
Student talks about debris.	6%	0%	0%	4%	1%	0%	0%	1%
Student says the Big Bang would not have been visible to the naked eye.	3%	0%	0%	2%	1%	17%	0%	5%
Student says the Big Bang would have been too hard to find.	0%	0%	0%	0%	1%	0%	0%	1%
Student says the Big Bang would have taken too long for someone to see.	4%	0%	9%	3%	1%	0%	0%	1%
Student says the Big Bang would have been too far away for someone to see.	4%	0%	5%	3%	0%	0%	0%	0%
Students says there was originally just energy.	0%	0%	0%	0%	16%	16%	0%	14%
Student says you could have seen the Big Bang from somewhere in the Solar System.	4%	0%	0%	3%	0%	0%	0%	0%
Student says you could see the Big Bang from elsewhere in the galaxy/another solar system or star.	9%	0%	0%	7%	0%	0%	0%	0%
Student says you could see the Big Bang from outside the Milky Way/from another galaxy.	4%	5%	5%	4%	0%	0%	0%	0%
Student says you could see the Big Bang from some not otherwise specified location.	4%	10%	0%	6%	0%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	9%	20%	5%	11%	4%	0%	7%	3%
Student says s/he has no idea.	6%	0%	5%	4%	0%	0%	0%	0%
Answer field is blank or the student provided no reason.	0%	5%	0%	1%	7%	0%	7%	5%

response:

> "No, There [*sic*] was nothing until the big bang so there was nowhere to watch from."

Is this student claiming that there was nothing - including time and space - until after the Big Bang? Or does her response indicate that she thinks there was no place to stand and watch the Big Bang since there was only empty space surrounding its location? This item, as written, had trouble eliciting responses that were sufficient to answer these questions and adequately categorize students' thoughts.

Likewise, Table 6.17 shows that many students answered a different question than we intended for item 4. We wanted students to state whether or not they think the universe has a center. A significant fraction of students either did not answer this question or did not provide reasons for their answers. Of the students who did answer the item and provided reasons for their answers, many did not say whether or not they think the universe has a center. Instead, they said if the universe had a center we would not be able to tell, either due to our ignorance or to our technological limitations. We do see evidence of larger shifts toward the correct reasons (which are the first two listed in Table 6.17) for the lecture-tutorial students compared to the non-lecture-tutorial students, but the overall pattern of responses highlighted this item as a candidate for revision.

Students' responses to items 5 and 6 lacked sufficient diversity to warrant creating and interpreting tables of reasoning elements as we did for items 1-4. Each student only received an overall score for her response. Table 6.13 shows that both lecture-tutorial and non-lecture-tutorial students move toward higher scores pre- to post-instruction. Note that students receive a score of 1 on item 5 when they talk about the distribution of galaxies petering out at some point in space. Table 6.13 also shows that a higher percent of students in the non-lecture-tutorial class earned a score of 2 on item 6 than students in the lecture-tutorial classes. Item 6 also stood out in our CTT analysis of Form B (see Section 6.3). How can we explain this? One possible explanation is that the subject of item 6 (specifically, that expansion does not affect distances between stars in a galaxy and planets in a solar system) is de-emphasized in our suite of cosmology lecture-tutorials compared to the

Table 6.17: Common reasoning elements used by students in their answers to Form B, Item 4: If you could travel to any location in the universe, could you ever see the center of the universe?

	Preinstruction				Postinstruction			
Reasoning Element	Class A	Class B	Class C	LT Classes	Class A	Class B	Class C	LT Classes
Student talks about how the universe is infinite and/or has no edges.	10%	15%	18%	11%	23%	11%	7%	20%
Student reasons using the idea that the universe looks the same no matter where you are.	0%	0%	4%	0%	11%	28%	7%	15%
Student says the Sun is in the center of the universe.	9%	0%	0%	7%	0%	0%	0%	0%
Student says the center is where the Big Bang happened/where everything is moving away from.	3%	20%	5%	7%	2%	0%	0%	1%
Student says our ignorance about the universe prevents us from seeing its center.	36%	15%	14%	31%	7%	6%	7%	7%
Student says the limitations of our technology prevent us from seeing its center.	6%	0%	4%	4%	4%	0%	0%	3%
Student says there is no center or the center changes because the universe is expanding and/or things are in motion.	16%	10%	18%	14%	7%	6%	7%	7%
Student says the universe has no center because it does not have a shape/has an irregular shape.	1%	0%	0%	1%	2%	0%	0%	1%
Student gives irrelevant information.	1%	5%	0%	2%	0%	0%	0%	0%
Student gives other reason not specified above.	9%	5%	4%	8%	14%	17%	29%	15%
Student says s/he has no idea.	4%	10%	0%	6%	0%	0%	0%	0%
Answer field left blank or no reason given for answer.	14%	25%	32%	17%	36%	33%	43%	35%

attention it received in Class C. Furthermore, we noticed a subset of students who did not simply circle one or more of the choices in Item 6. For example, some circled the words “always increase” and crossed out the words “never decrease” for each of the three choices. These responses indicated that item 6 may have been misinterpreted by some students and requires revision.

6.5.3 Form C Responses

Table 6.18 shows the distribution of scores on Form C for Classes A-C as well as for the two lecture-tutorial classes combined. Even more than Form B, the lecture-tutorial students show larger improvements in their scores on Form C than their non-lecture-tutorial peers.

Figure 6.16 shows the percent of students in each class, as well as for the lecture-tutorial group overall, that selected choices A-D on item 1. The correct answer is C. Figure 6.16 shows that a larger percent of lecture-tutorial students selected C post-instruction than non-lecture-tutorial students. Furthermore, the change in the percentage of students choosing C pre- to post-instruction is greater for lecture-tutorial students than non-lecture-tutorial students (in fact, the shift is negative for the non-lecture-tutorial students).

Table 6.19 shows the reasons students gave for their answers pre-instruction and Table 6.20 shows the reasons students gave post-instruction. Tables 6.19 and 6.20 also give the percent of students using a particular reasoning element for each of the four answer choices (A-D). What patterns emerge from this data?

Pre-instruction, some students think of light-years as a unit of time. Some students claim light-years are larger than regular years, some claim they are shorter than regular years, and some claim they are different than regular years but do not specify if they think light-years are larger or shorter. These non-expert-like ideas about light-years do not appear in any of the post-instruction responses.

Post-instruction, the majority of students who select option B do so because they claim light must take eight billion years to traverse eight billion light-years. While this response is consistent with how many Astro 101 students are exposed to the ideas of light travel time and lookback time,

Table 6.18: Pre - and Post-instruction distribution of scores on Form C for fall 2009.

		Class A				Class B				Class C				LT Classes			
Preinstruction	Item	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
	Item 1	0%	67%	14%	19%	0%	71%	0%	29%	3%	40%	3%	53%	0%	68%	11%	21%
	Item 3	6%	81%	6%	7%	21%	54%	0%	25%	13%	43%	23%	20%	10%	74%	4%	12%
	Item 4	4%	71%	24%	0%	4%	58%	38%	0%	10%	60%	27%	3%	4%	68%	28%	0%
	Item 5	11%	69%	4%	16%	13%	58%	8%	21%	13%	40%	10%	37%	12%	66%	5%	17%
Postinstruction		Class A				Class B				Class C				LT Classes			
	Item	0	1	2	3	0	1	2	3	0	1	2	3	0	1	2	3
	Item 1	0%	11%	8%	80%	5%	37%	0%	58%	4%	43%	0%	54%	1%	18%	6%	75%
	Item 3	0%	10%	21%	69%	5%	11%	42%	42%	11%	14%	32%	43%	1%	10%	26%	63%
	Item 4	3%	26%	13%	57%	11%	63%	16%	11%	14%	46%	25%	14%	5%	35%	14%	46%
	Item 5	3%	16%	21%	59%	5%	37%	16%	42%	21%	39%	7%	32%	4%	21%	20%	55%

Table 6.19: Common pre-instructional reasoning elements used by students in their answers to Form C, Item 1: Galaxy X and Galaxy Y are currently 8 billion light-years apart in the expanding universe. How long will light from Galaxy X take to reach Galaxy Y?

	Class A				Class B				Class C				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther part.	0%	3%	65%	8%	0%	0%	100%	33%	0%	0%	94%	67%	0%	2%	73%	13%
Student talks about galaxies getting closer together.	0%	0%	0%	17%	0%	0%	0%	33%	100%	0%	0%	33%	0%	0%	0%	20%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	0%	100%	0%	86%	0%	0%	0%	0%	100%	0%	33%	0%	96%	0%	7%
Student says light-years are shorter than normal years.	33%	0%	0%	9%	0%	0%	0%	0%	0%	0%	0%	0%	33%	0%	7%	0%
Student says light-years are longer than normal years.	0%	0%	13%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	10%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	9%	17%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	7%	13%
Student says light may be slowed by something between the two galaxies.	0%	0%	4%	8%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	7%
Student says there's not enough information.	0%	0%	0%	92%	0%	0%	0%	67%	0%	0%	0%	67%	0%	0%	0%	87%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	67%	0%	4%	0%	0%	0%	0%	0%	0%	0%	6%	33%	67%	0%	3%	0%
Student says s/he has no idea.	0%	0%	0%	8%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	7%
Response field left blank or no reason given.	0%	0%	4%	0%	0%	14%	0%	33%	0%	0%	0%	0%	0%	4%	3%	7%

Table 6.20: Common post-instructional reasoning elements used by students in their answers to Form C, Item 1: Galaxy X and Galaxy Y are currently 8 billion light-years apart in the expanding universe. How long will light from Galaxy X take to reach Galaxy Y?

	Class A				Class B				Class C				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther part.	0%	0%	91%	0%	0%	0%	100%	0%	100%	0%	100%	0%	0%	0%	92%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	60%	0%	0%	0%	67%	0%	0%	0%	80%	0%	0%	0%	64%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	0%	0%	0%	0%	100%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	100%	0%	7%	0%	0%	0%	0%	0%	0%	0%	0%	0%	100%	0%	6%	0%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Response field left blank or no reason given.	0%	40%	2%	0%	0%	33%	0%	0%	0%	20%	0%	100%	0%	36%	2%	0%

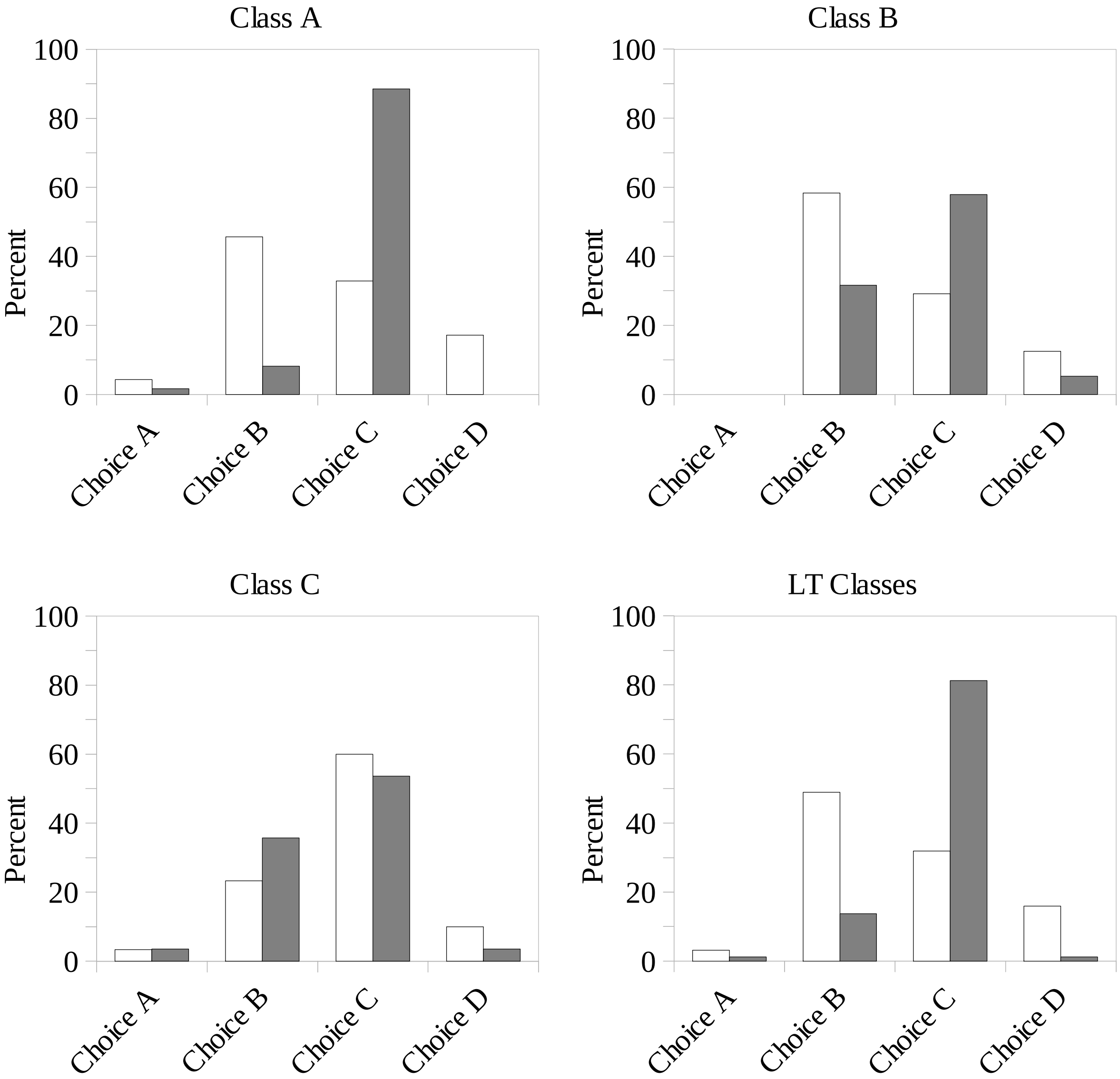


Figure 6.16: Percent of students who selected choice A, B, C, or D in response to item 1 on Form C. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

it ignores the effects of the expansion of the universe, as discussed in Chapter 3.

The expansion of the universe is the most common reason students cite for choosing option C. The percent of students who defend their choice using the idea that the universe is expanding increases for all groups. Note, however, that the percent of students in each score category for item is relatively unchanged for Class C, the non-lecture-tutorial class (see Table 6.18). This is in contrast to the lecture-tutorial classes, which exhibit significant improvements in their overall scores to item 1 pre- to post-instruction. This shows that while both lecture-tutorial and non-lecture-tutorial students can choose the correct answer for the correct reason, the lecture-tutorial students are much more likely to change their answer to C pre- to post-instruction and for the correct reason.

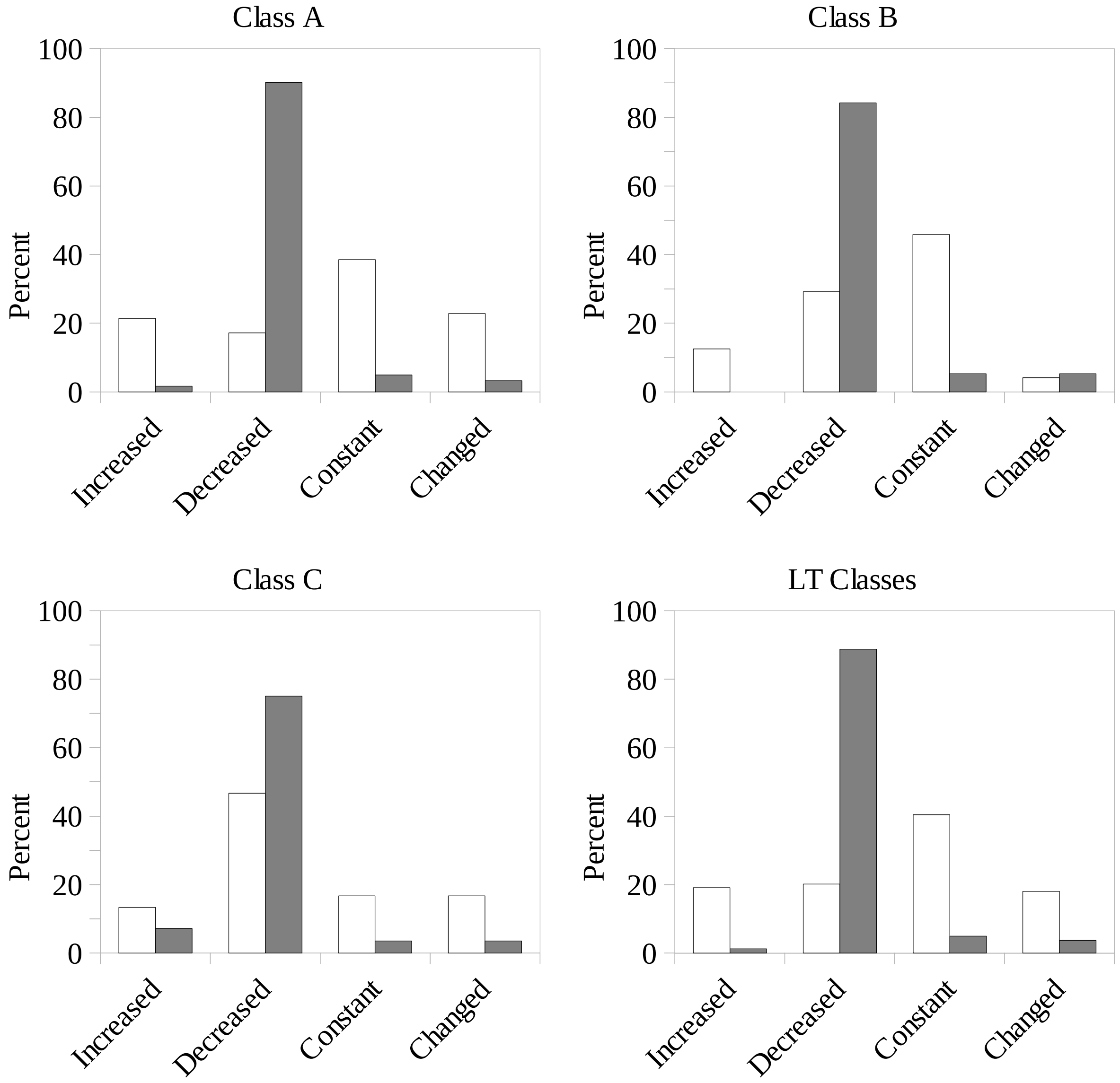


Figure 6.17: Percent of students who think the temperature of the universe has increased, decreased, stayed constant, or changed (without specifying if it went up or down). White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Figure 6.17 shows the percent of students who say the temperature of the universe has increased, decreased, remained constant, or changed (but without specifying if it went up or down). Tables 6.21 and 6.22 show the percent of students who used each reasoning element for each of the four responses plotted in Figure 6.17 (I = the temperature increased, D = the temperature decreased, S = the temperature stayed the same, and C = the temperature changed).

As Figure 6.17 shows, all students, regardless of whether they used the lecture-tutorials or not, move toward saying that the temperature of the universe has decreased over time. However, the change is greater for the lecture-tutorial students than for the non-lecture-tutorial students. Compared to the non-lecture-tutorial students, a smaller percent of the lecture-tutorial students pre-instruction and a greater percent of the lecture-tutorial students post-instruction said that the temperature is decreasing.

We argue that the reasoning elements in Tables 6.21 and 6.22 as well as the distribution of responses in Figure 6.17 are consistent with the idea that students probably do not have any robust misconceptions about the temperature of the universe. Rather, they are constructing their answers using their cognitive resources, phenomenological experiences, and pre-existing knowledge. For example, many students talk about the birth, death, and changes that occur during the lifetimes of celestial objects. One student who argued that the temperature of the universe is changing framed her response in terms of her experiences here on Earth:

> "The temperature on Earth changes due to human activity as well as natural phenomena like volcanoes that emit elements. I think other planets do the same w/their natural elements, so with all the planets changing, I think the universe had to have had a slight alteration in temperature."

Another argued that the temperature must be constant due to competing effects canceling out:

> "I would say it has probably remained about the same since stars die and are born."

These responses indicate that students are drawing upon their pre-existing knowledge to answer this item.

Table 6.23 shows the most common responses to item 4, pre- and post-instruction. Pre-instruction, over half (53% of the non-lecture-tutorial and 56% of the lecture-tutorial students)

Table 6.21: Common pre-instructional reasoning elements used by students in their answers to Form C, Item 3: Has the temperature of the universe changed over time, or has it always been about the same?

	Class A				Class B				Class C				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	13%	50%	4%	0%	0%	100%	27%	0%	25%	43%	0%	0%	11%	68%	11%	0%
Student talks about the birth/formation of stars/a star.	13%	0%	11%	19%	33%	0%	27%	0%	25%	0%	0%	20%	17%	0%	16%	18%
Student talks about the birth/formation of planets/a planet.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of galaxies/a galaxy.	13%	8%	0%	6%	33%	0%	0%	0%	25%	0%	0%	0%	17%	5%	0%	6%
Student talks about the birth/formation of unspecified objects.	0%	0%	4%	0%	0%	0%	9%	0%	0%	7%	0%	0%	0%	0%	5%	0%
Student talks about the death of stars.	0%	0%	7%	13%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	12%
Student talks about the death of planets.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%
Student talks about changes during the lives of stars/a star.	13%	33%	0%	6%	0%	0%	0%	0%	25%	0%	20%	20%	11%	21%	0%	6%
Student talks about changes during the lives of planets/a planet.	40%	17%	26%	38%	0%	0%	0%	0%	0%	0%	0%	0%	33%	11%	18%	35%
Student talks about changes during the lives of galaxies/a galaxy.	0%	0%	0%	19%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	18%
Student talks about changes during the lives of unspecified objects.	0%	0%	4%	6%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	6%
Student talks about how the universe is big.	0%	8%	11%	0%	0%	0%	0%	0%	0%	0%	20%	0%	0%	5%	8%	0%
Student talks about competing effects canceling out.	7%	0%	22%	0%	0%	0%	9%	0%	0%	0%	0%	0%	6%	0%	18%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	8%	0%	6%	33%	0%	9%	100%	0%	21%	0%	0%	6%	5%	3%	12%
Student says s/he has no idea.	7%	8%	11%	0%	33%	14%	18%	0%	0%	0%	0%	0%	11%	11%	13%	0%
Response field left blank or no reason given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 6.22: Common post-instructional reasoning elements used by students in their answers to Form C, Item 3: Has the temperature of the universe changed over time, or has it always been about the same?

	Class A				Class B				Class C				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	0%	76%	33%	50%	0%	50%	0%	100%	0%	57%	0%	0%	0%	70%	25%	67%
Student talks about the birth/formation of stars/a star.	0%	0%	0%	50%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	33%
Student talks about the birth/formation of planets/a planet.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of galaxies/a galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of unspecified objects.	0%	0%	33%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	25%	0%
Student talks about the death of stars.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of planets.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of stars/a star.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of planets/a planet.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of galaxies/a galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of unspecified objects.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about how the universe is big.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about competing effects canceling out.	0%	0%	33%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	25%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	4%	0%	0%	0%	0%	0%	0%	0%	5%	100%	0%	0%	3%	0%	0%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%
Response field left blank or no reason given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 6.23: Common reasoning elements used by students in their answers to Form C, Item 4: How does the total amount of matter in the universe right now compare to the total amount of matter in the universe at the very beginning of the universe (the moment just after the Big Bang)?

	Pre-instruction				Post-instruction			
Reasoning Element	Class A	Class B	Class C	LT Classes	Class A	Class B	Class C	LT Classes
Student says there was no matter/only energy in the beginning.	0%	0%	0%	0%	46%	0%	11%	35%
Student talks about the temperature cooling and/or that matter formed from energy.	0%	0%	3%	0%	21%	11%	4%	19%
Student says there is more matter now because the universe is expanding.	4%	17%	10%	7%	3%	0%	7%	3%
Student says the amount of matter increases as objects form and/or evolve.	17%	17%	20%	17%	3%	11%	7%	5%
Student says the amount of matter decreases as objects form and/or evolve.	9%	4%	10%	7%	0%	0%	11%	0%
Student says the amount of matter increases as objects interact and/or die.	1%	13%	7%	4%	0%	0%	0%	0%
Student says the amount of matter decreases as objects interact and/or die.	3%	4%	0%	3%	0%	6%	0%	0%
Student says the amount of matter does not change.	59%	50%	53%	56%	23%	63%	36%	33%
Student gives irrelevant information.	0%	0%	0%	0%	2%	0%	0%	1%
Student gives some other reason not specified above.	3%	4%	3%	3%	3%	5%	4%	4%
Student says s/he has no idea.	4%	0%	1%	3%	0%	1%	0%	0%
Response field left blank and/or no response given.	4%	4%	10%	4%	8%	0%	21%	9%

say that the total amount of matter has not changed over time. A smaller percentage (20% of the non-lecture-tutorial and 17% of the lecture-tutorial students) say that the amount of matter increased as objects in the universe formed and/or evolved. For example, one student wrote

> "There is more matter now b/c after the Big Bang there were only rocks and cells. There wasn't [*sic*] any plants or animals or people."

Almost none of the students gave a correct answer (one or both of the first two responses listed in Table 6.23 pre-instruction).

The post-instruction results for the lecture-tutorial and non-lecture-tutorial students show significant differences. 35% of the lecture-tutorial students said there was no matter/only energy at the very beginning of the universe and 19% talked about the temperature of the universe cooling and matter forming from energy. For the non-lecture-tutorial classes, these percentages are 11% and 4%, respectively. The percent of students claiming the total amount of matter remained constant dropped for both populations, to 36% for the non-lecture-tutorial students and 33% for the lecture-tutorial students. The percent of students saying the amount of matter increases as objects form and/or evolve also decreased for both populations. We conclude that while instruction helped all students, the lecture-tutorial students performed better than the non-lecture-tutorial students on this item post-instruction.

Figure 6.18 shows the percent of students who say the density of matter in the universe has increased, decreased, remained constant, or changed (but without specifying if it went up or down). There is a definite movement of students in the lecture-tutorial classes post-instruction toward the correct answer (decreased). The same cannot be said for the non-lecture-tutorial students, for which the percent saying "decreased" actually declined pre- to post-instruction.

What are students actually saying in their responses? Table 6.24 shows the reasoning elements used by students pre-instruction and Table 6.25 shows the reasoning elements used by students post-instruction. Each table shows the percent of students who used a particular reasoning element for each of the four possible answers (I = increased, D = decreased, S = stayed the same, and C = changed). Among students who said the density has decreased, the most popular reasoning element

Table 6.24: Common pre-instructional reasoning elements used by students in their answers to Form C, Item 5: Has the density of matter in the universe changed over time, or has it always been about the same?

	Class A				Class B				Class C				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C	I	D	S	C
Student says the amount of matter in the universe has changed.	6%	7%	6%	13%	17%	14%	0%	0%	0%	0%	0%	0%	9%	10%	5%	11%
Student says the amount of matter in the universe has not changed.	13%	0%	13%	25%	0%	0%	0%	0%	0%	13%	0%	0%	9%	0%	9%	22%
Student says the size of the universe has changed.	0%	86%	13%	6%	33%	86%	0%	0%	0%	75%	50%	0%	9%	86%	9%	6%
Student says the size of the universe has not changed.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time	63%	7%	6%	31%	33%	0%	0%	100%	50%	0%	0%	33%	55%	5%	5%	39%
Student talks about objects evolving over time.	13%	0%	0%	0%	33%	0%	0%	0%	25%	6%	0%	33%	18%	0%	0%	0%
Student talks about objects dying/breaking up.	6%	0%	0%	6%	17%	0%	0%	0%	0%	0%	0%	33%	9%	0%	0%	6%
Student talks about the effects of gravity.	13%	7%	6%	0%	0%	0%	0%	0%	0%	0%	0%	0%	9%	5%	5%	0%
Student talks about matter changing forms.	6%	0%	0%	19%	0%	0%	0%	0%	0%	0%	0%	33%	5%	0%	0%	17%
Student gives some other reason not specified above.	0%	0%	6%	0%	33%	0%	17%	0%	13%	0%	0%	0%	9%	0%	9%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says s/he has no idea.	6%	0%	6%	6%	0%	0%	0%	0%	13%	13%	50%	0%	5%	0%	5%	6%
Response field is blank and/or no response given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 6.25: Common post-instructional reasoning elements used by students in their answers to Form C, Item 5: Has the density of matter in the universe changed over time, or has it always been about the same?

	Class A				Class B				Class C				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C	I	D	S	C
Student says the amount of matter in the universe has changed.	0%	0%	20%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	13%	0%
Student says the amount of matter in the universe has not changed.	0%	0%	0%	0%	0%	9%	0%	0%	0%	9%	0%	0%	0%	2%	0%	0%
Student says the size of the universe has changed.	0%	73%	0%	0%	0%	73%	33%	0%	50%	82%	0%	25%	0%	73%	13%	0%
Student says the size of the universe has not changed.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time	0%	0%	0%	0%	33%	0%	0%	100%	50%	0%	0%	0%	25%	0%	0%	20%
Student talks about objects evolving over time.	0%	0%	20%	0%	0%	0%	0%	0%	0%	0%	0%	50%	0%	0%	13%	0%
Student talks about objects dying/breaking up.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the effects of gravity.	0%	0%	0%	0%	33%	0%	0%	0%	50%	0%	0%	0%	25%	0%	0%	0%
Student talks about matter changing forms.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	0%	4%	0%	25%	0%	0%	0%	0%	0%	0%	0%	0%	0%	3%	0%	20%
Student gives irrelevant information.	0%	2%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	2%	0%	0%
Student says s/he has no idea.	0%	0%	20%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	13%	0%
Response field is blank and/or no response given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

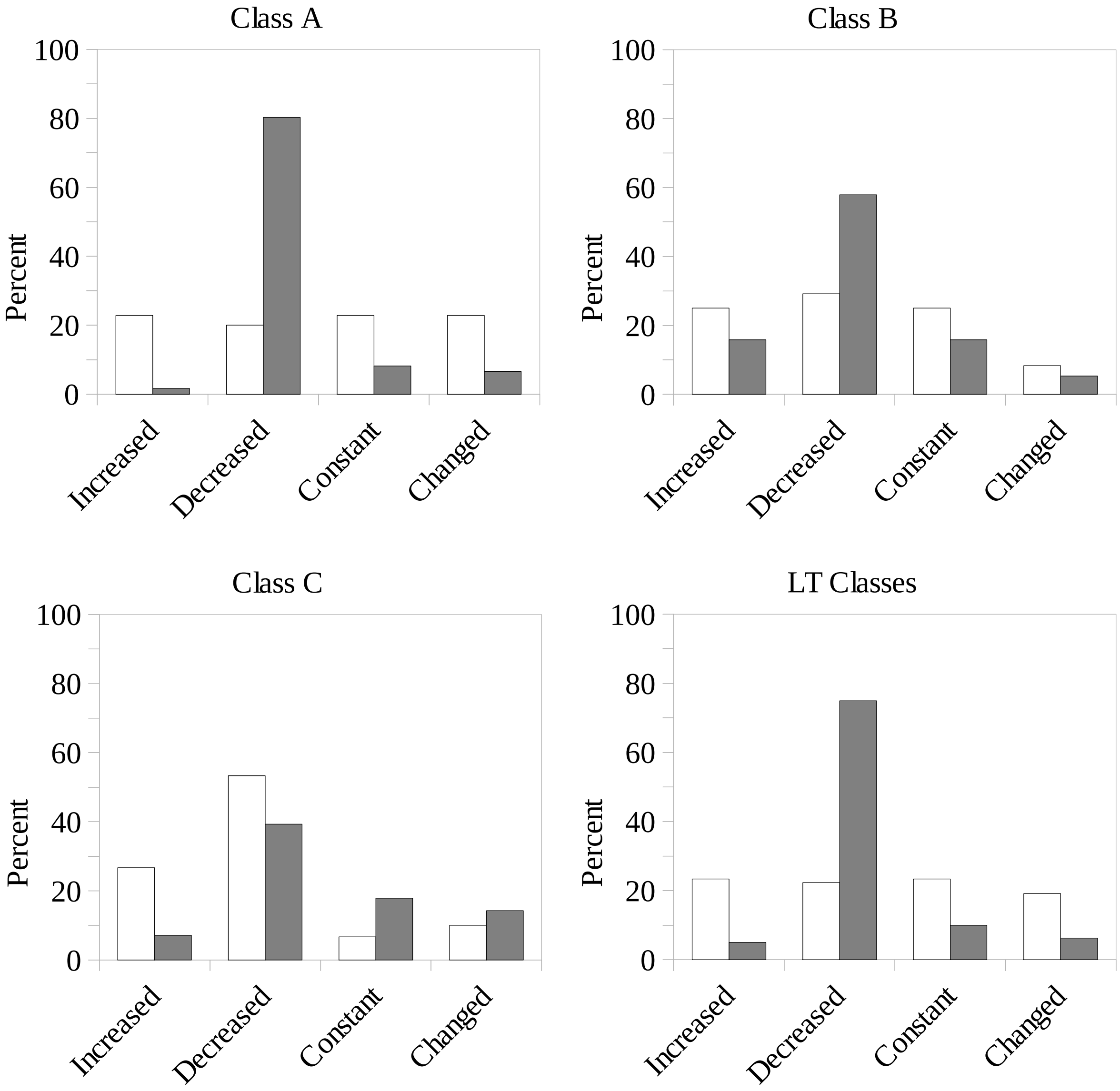


Figure 6.18: Percent of students who think the density of matter in the universe has increased, decreased, stayed constant, or changed (without specifying if it went up or down). White bars represent pre-instruction responses and grey bars represent post-instruction responses.

is "the size of the universe has changed," which was invoked by 86% (73%) of the lecture-tutorial students and 75% (82%) of the non-lecture-tutorial students pre-instruction (post-instruction). The most popular reasoning element used by students who gave a wrong answer is "objects form over time." For example, one student wrote

> "I believe there is more now because more things have been created."

Table 6.26: Pre- and Post-instruction distribution of scores on Form D for fall 2009.

		Class A			Class B			Class C			LT Classes		
Pre-instruction	Item	0	1	2	0	1	2	0	1	2	0	1	2
	Item 1	1%	80%	19%	0%	91%	9%	0%	93%	7%	1%	84%	15%
	Item 2	3%	28%	69%	0%	16%	84%	3%	10%	86%	2%	24%	75%
	Item 4	7%	38%	55%	5%	25%	70%	13%	24%	62%	6%	33%	61%
Post-instruction		Class A			Class B			Class C			LT Classes		
	Item	0	1	2	0	1	2	0	1	2	0	1	2
	Item 1	0%	13%	87%	3%	32%	65%	0%	79%	21%	1%	20%	79%
	Item 2	2%	13%	85%	3%	21%	76%	0%	16%	84%	2%	16%	82%
	Item 4	8%	43%	48%	3%	18%	79%	11%	47%	42%	6%	34%	60%

This reasoning element was much less popular in the post-instruction results for the lecture-tutorial students than for the non-lecture-tutorial students. This further supports the hypothesis that the lecture-tutorials help students on the *evolving universe* construct.

6.5.4 Form D Responses

Table 6.26 shows the distribution of scores on Form D for the fall 2009 semester. Overall, the lecture-tutorial students show a greater improvement pre- to post-instruction than the non-lecture-tutorial students. Since we did not delineate any specific reasoning elements on these items (see the scoring rubric in Appendix C), we cannot breakdown students's responses any further, except for item 1. Figure 6.19 shows students' graph selections for item 1. Note that a much greater percent of lecture-tutorial students select the correct rotation curve for a spiral galaxy (graph 2) post-instruction than their non-lecture-tutorial peers. This result is striking because item 1 itself could simply be considered a factual recall question; our data, however, suggest that going through the lecture-tutorial activity (in which students must articulate for themselves why flat rotation curves are evidence for dark matter) may significantly help their abilities to remember the correct answer.

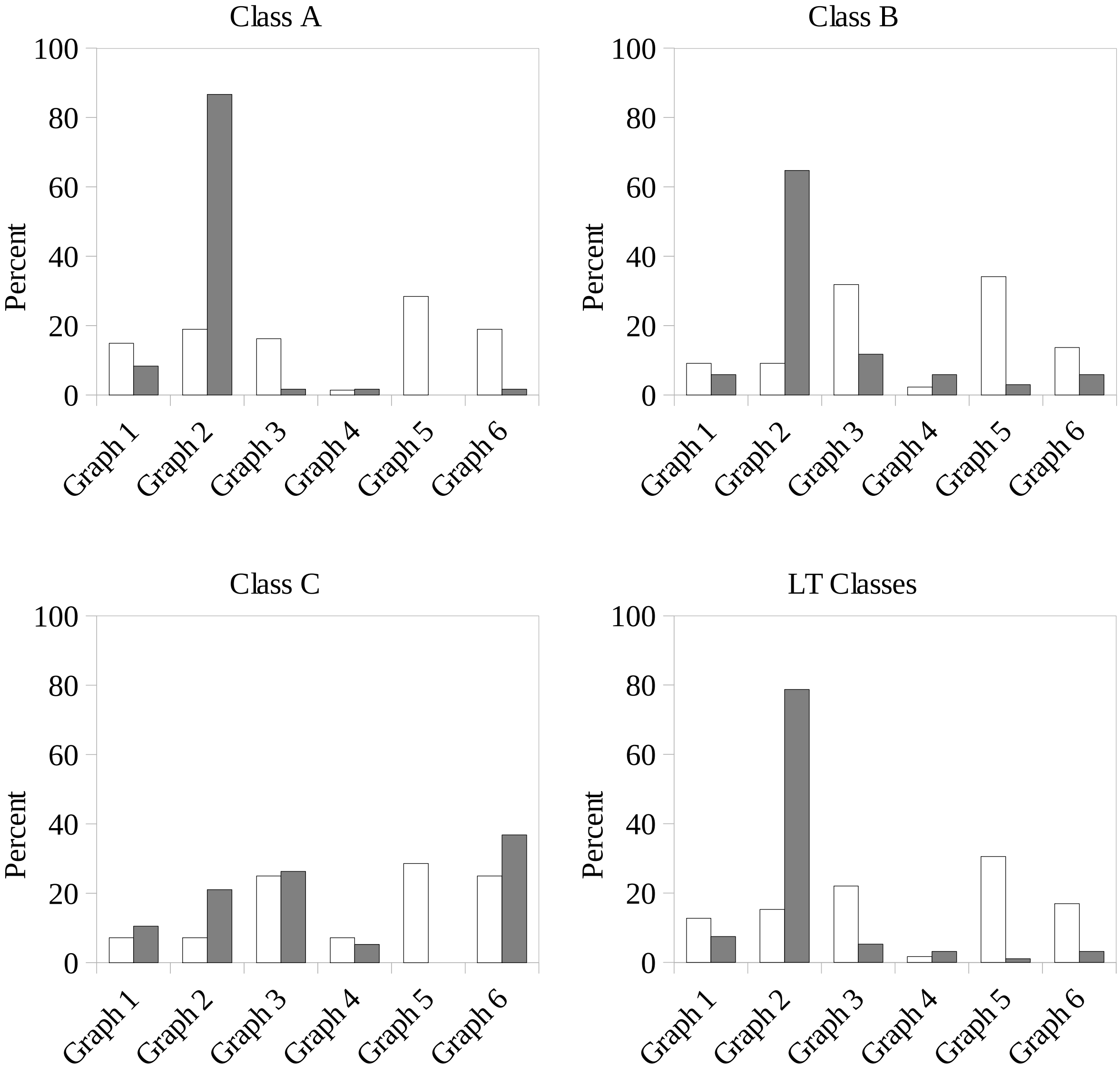


Figure 6.19: Students' graph choices for item 1 on Form D. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

6.6 Validity

In Sections 6.3.3 and 6.4.5 we looked at whether or not our surveys are reliable – or, in other words, whether or not they give us consistent and accurate results. But a more fundamental question is "Do the surveys actually measure what we intend them to measure?" Answering this question falls into the domain of assessing the *validity* of the surveys, which is the subject of this section.

Validity is not a property of a test *per se*. Rather, it is associated with the interpretation one gives to test scores (AERA, APA, and NCME 1999; Kane 1992; Wilson 2005). Kane summarizes this view of validity by explaining how validity involves an interpretive argument:

> "A test-score interpretation always involves an *interpretive argument*, with the test score as a premise and the statements and decisions involved in the interpretation as conclusions. The inferences in the interpretive argument depend on various assumptions, which may be more-or-less credible. [...] Because it is not possible to prove all of the assumptions in the interpretive argument, it is not possible to verify this interpretive argument in any absolute sense. The best that can be done is to show that the interpretive argument is highly plausible, given all available evidence" (p. 527, italics in original).

Kane further recommends that test designers explicitly state the assumptions for which they must find evidence to support their interpretations of and decisions based on the test's scores (Kane 1992).

What interpretations and decisions do we want to make? We want to know if students learn more cosmology if they use the cosmology lecture-tutorials. We want to interpret survey scores such that higher scores imply a greater mastery of the relevant construct. We want our comparisons of pre-instructional scores to post-instructional scores to tell us what effect, if any, instruction had on students' cosmology knowledge. We want to compare the gains of students who used the lecture-tutorials to the gains of students who did not to see if the lecture-tutorials had any effect. These interpretations rest on the following assumptions:

(1) Each survey adequately covers the construct it is intended to measure.

(2) The students who take the surveys are representative of the target population of Astro 101 students – that is, we can generalize our results.

(3) Astro 101 students correctly read and interpret our survey items.

(4) Students' responses reveal their ideas about cosmology.

(5) Students' responses can be reliably transformed into numerical scores.

(6) These scores can be used to find measurable differences between different populations of students.

(7) Differences in pre- and post-instruction scores are due to the instruction the students received.

(8) Differences in the learning gains of students who have and have not used the lecture-tutorials are due to the lecture-tutorials and not some other variable.

This is a large list and we cannot support every entry with the same amount of evidence. To reiterate Kane (1992), we cannot provide any absolute proof of the validity of our surveys. Furthermore, we will shore up some weaknesses in our validity evidence when we present the data we collected from subsequent semesters (Chapters 7 and 8). However, at this point we are able to begin addressing each of these assumptions as part of our validity argument.

Each survey adequately covers the construct it is intended to measure. We devoted much of the survey design process (Chapter 4) to addressing this issue. In Chapter 4, I described how we defined each construct (drawing on the research literature where available), developed our construct maps, and designed our surveys' items. Appendix C also contains our detailed scoring rubrics for each item, which demonstrates the level of detail we expect from students as part of a correct answer. Finally, our items were reviewed by several experts in the fields of astrophysics and education. We conducted this careful and detailed design and development process in order to support, as recommended by the AERA, APA, and NCME (1999), the "relationship between [each] test's content and the construct it is intended to measure."

The students who take the surveys are representative of the target population of Astro 101 students - that is, we can generalize our results. We surveyed a large number of students (501 pre-instruction and 406 post-instruction, according to Table 6.1) because we wanted results that are robust and generalizable. However, all of the students in our fall 2009 sample came from large, research-focused, flagship institutions of states in the western United States. We will therefore withhold any further comment on the generalizability of our results until we discuss our data from the fall 2010 semester (Chapter 8), which includes students from a variety of institution types spread across the country.

Astro 101 students correctly read and interpret our survey items. To help address this issue, I interviewed four Astro 101 students during the fall 2009 semester. All four students volunteered for an interview by checking a box on the consent forms the signed when they agreed to participate in this study. They received no compensation for their time.

I interviewed these four students in order to get a sense of whether or not Astro 101 students were interpreting our survey items as we intended. The interviews were semi-structured think-alouds (Otero and Harlow 2009): I gave each student one survey item at a time and asked him/her to describe everything he/she thought of while constructing his/her answer (Willis 2005). Before the students tried thinking aloud on the survey items, I gave them an unrelated question on which to practice thinking aloud (Otero and Harlow 2009; Willis 2005). I also followed much of Patton's (1980) advice: I avoided dichotomous and "why" questions, I had one idea per question, and I made sure that the student did the lion's share of the talking during the interview. I prepared several follow-up questions to each item, in case a student did not adequately explain a key point of interest (Willis 2005).

I videotaped all four interviews. I also took notes during the interviews and recorded my summary and impressions of each interview immediately after the interviewee left (Erickson 1986; Patton 1980).

I will hereafter refer to the four interviewed students by the pseudonyms Melissa, Tyson, Kelsey, and Abigail. Table 6.27 breaks down which surveys were used during each interview. For example, I first interviewed Melissa on the items on Form B, followed by Form A, and ending with Form D. I varied the order in which students covered each form because people tend to give poorer quality of responses to questions asked early in an interview (Willis 2005).

Overall, the four students appeared to interpret the items as we intended, with a few exceptions. For example, on Form D both Melissa and Kelsey interpreted the term "rotation curve" to be synonymous with "orbit," although this did not appear to affect how they answered Form

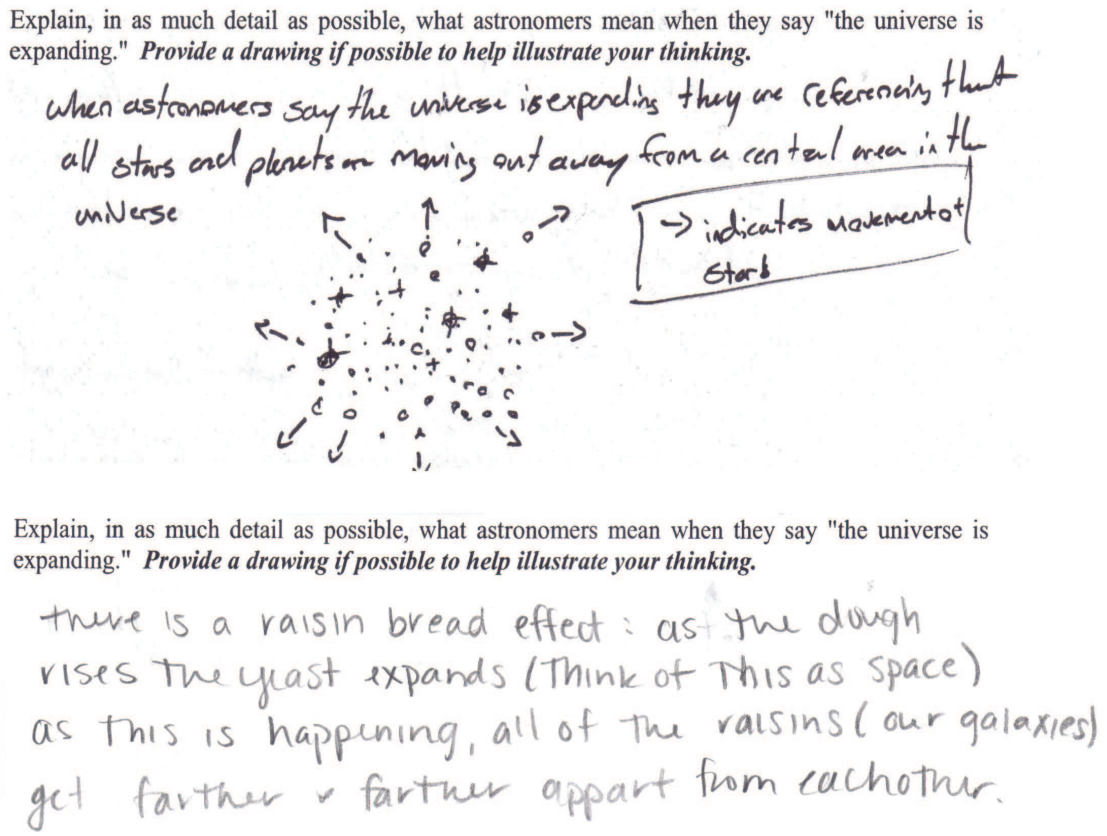
Table 6.27: The forms each student responded to during his/her interview in the fall 2009. The numbers denote the order in which I presented the surveys to the student.

Student	Form A	Form B	Form C	Form D
Melissa	2	1	-	3
Tyson	3	2	1	-
Kelsey	-	3	2	1
Abigail	1	2	3	-

D's items. Melissa, Tyson, and Abigail were also confused by the wording of item 5 on Form A. Tyson, for example, interpreted "old" to mean "what the universe was like in the past." All of the students struggled to understand what item 2 on Form C was asking (as noted above, we ultimately eliminated this item from our analysis). Finally, all the students noted that we needed a "none of the above" option for some of the items in which students must choose an answer. The remaining items did not seem to pose serious issues for these four students.

In addition to the items highlighted by Melissa, Tyson, Kelsey, and Abigail, our analysis of the written responses to our survey questions identified some potentially problematic items. These include item 5 on Form A, items 3, 4, and 6 on Form B, and items 3, 4, and 5 on Form C. All of these items were highlighted by some aspect of our analysis described above, and all of these items were revised for subsequent versions of the surveys.

Students' responses reveal their ideas about cosmology. As the previous section illustrates, our survey items do elicit a wide range of student ideas about the Big Bang, the expansion and evolution of the universe, and dark matter. For example, Figure 6.20 shows the responses of two different students to the same item. This one item was able to elicit responses that indicate very different conceptions of the expansion of the universe from these students. Our CTT and IRT analyses show that our items vary in terms of their difficulties and generally span the range of student abilities. The outfit statistics for respondents discussed in Section 6.4.4 indicate that students' responses are close to their theoretical expectations. While some revision of the items was necessary, as discussed above and in Section 6.7 below, our items for the fall 2009 semester provided a respectable starting point for this study.



Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." ***Provide a drawing if possible to help illustrate your thinking.***

when astronomers say the universe is expanding they are referencing that all stars and planets are moving out away from a central area in the universe

→ indicates movement of stars

Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." ***Provide a drawing if possible to help illustrate your thinking.***

there is a raisin bread effect : as the dough rises the yeast expands (Think of This as space) as this is happening, all of the raisins (our galaxies) get farther & farther appart from eachother.

Figure 6.20: Two different responses from two different students to the same item on Form B.

Students' responses can be reliably transformed into numerical scores. We addressed the issue of reliability using Cronbach's α (Section 6.3.3), item response theory (Section 6.4.5), and by looking at the ability of multiple raters to arrive at the same scores. While our calculated values for Cronbach's α are lower than we would like, our IRT and inter-rater reliability analyses indicate that our surveys are, in general, reliable.

These scores can be used to find measurable differences between different populations of students. Much of our quantitative analysis presented in this chapter supports this point. Except for Form A, we found significant and measurable differences between students pre- and post-instruction and between students that did and did not use the lecture-tutorials using classical test theory, item response theory, and even just by examining the distribution of raw scores.

Differences in pre- and post-instruction scores are due to the instruction the students received. We cannot yet address this point because we did not make daily and detailed classroom observations during the fall of 2009. We made such observations during the spring and fall 2010 semesters. I will describe these observations in more detail in the sections of Chapters 7 and 8 that re-address these validity issues (Sections 7.5 and 8.5).

Differences in the learning gains of students who have and have not used the lecture-tutorials are due to the lecture-tutorials and not some other variable. As with the previous assumption, we postpone our discussion of this point until Chapters 7 and 8 since we only made classroom observations in the spring and fall of 2010.

6.7 Revisions

Sections 6.3-6.5 give detailed analyses of our data using classical test theory, item response theory, and by looking at students' written responses. At various points during these sections, I highlighted where the data implies revisions are necessary. This section outlines the revisions we made to both the surveys and the lecture-tutorials in preparation for the spring 2010 semester.

6.7.1 Surveys

Of all the conceptual surveys, Form A's items were flagged the most by our analyses. Some of these issues, such as the fact that the lecture-tutorial students did not show any significant improvement over their non-lecture-tutorial counterparts, suggest revisions for the lecture-tutorials. Others, however, concern Form A itself. Specifically, we found the following issues with Form A:

(1) The value of Cronbach's α could be improved if Form A was longer.

(2) The distribution of respondent outfit statistics was skewed toward slightly higher values than expected.

(3) The outfit statistic for item 1 was larger than expected.

(4) Form A's Wright map and standard error of measurement plot indicate that Form A needs more easier items in order to more accurately estimate abilities for students of low ability.

(5) The interviewed students found the wording of item 5 confusing.

(6) The interviewed students also requested the option of answering "none of the above" to many items.

We considered all these factors when we revised Form A.

Our revised version of Form A (along with revised versions of Forms B-D) is located in Appendix D. We added the following sentence to the end of items 1-4: "If your answer is 'none,' explain why." We also removed our original item 5 and added two new items in its place. The new item 5 and item 6 both refer to a Hubble plot for a universe expanding at a constant rate. Item 5 asks students to draw what this plot would look like if the universe expanded twice as fast. Item 6 asks students to draw what this plot would look like for a much older universe. Both of these items are meant to probe students' understandings of the relationships between the slope of a Hubble plot, the expansion rate of the universe, and the age of the universe. We also intended them to be easier than items 1-4.

We also revised the other three surveys. On Form B, our analyses also suggested that a longer test may improve Form B's Cronbach's α. We further found an unexpectedly high value for Yen's Q3 statistic for item 1 and item 5 and our analysis of students' written responses revealed that students may not have interpreted items 3, 4, and 6 as we intended. We therefore made several changes to Form B. First, we added a new item, which became item 3 of the revised survey (items 3-6 of the fall 2009 version became items 4-7 of the spring 2010 version). The new item 3 gives students the drawing of our observable universe from the beginning of the "Making Sense of the Universe and Expansion" lecture-tutorial and asks them to describe what inhabitants of a galaxy at the edge of our observable universe would see if they looked at regions beyond Earth's observable universe. We intended this item to probe whether or not students think there are galaxies beyond our observable universe. We also revised the fall 2009 survey's item 3 (now item 4) such that students have to select one of three possible choices that best describes how they think of the Big Bang. The fall 2009 survey's item 4 (now item 5) was rewritten to begin with the question "Independent of whether we know its true location, is there a center to the universe?" We made this revision in response to the number of students who gave responses such as "we would not be able to find the center of the universe," which did not provide us the information we were looking for. Finally, we also revised the wording and answer choices in the previous version's item 6 (now item 7) to clarify for students that we want them to circle one or more of the choices, not just parts of each choice.

On Form C, we moved items 3-5 to the beginning of the survey, making them items 1-3. We also tweaked the wording of the old item 3 (now item 1) and item 5 (now item 3) to prompt students to say whether or not the temperature and density have increased, decreased, or stayed the same; this was in response to the number of students who said these properties changed but did not say how. Items 4-6 of the new Form C are similar in content to item 1 of the fall 2009 version of Form C in that they all address the relationship between expansion, distances, light travel times, and lookback times. We decided to have multiple items focusing on these relationships in order to increase the overall number of items on Form C (and thus possibly raise its value of Cronbach's α). We also wanted to add more difficult items to Form C, which was an issue raised by Form C's Wright map.

Compared to the other three surveys, we made relatively few changes to Form D. Our changes were primarily concerned with improving the clarity of Form D's items. We changed the wording slightly on item 1 and explicitly asked students to provide rankings on item 2 and 3. Item 3 remained on Form D because we had not yet realized our error in the rankings of the net gravitational forces felt by objects of increasing radii in the galaxy (see Section 4.6). We revised item 4 by adding three choices from which students could chose. We did not make any further changes to Form D until after the spring 2010 semester.

6.7.2 Lecture-Tutorials

Appendix E contains our revised lecture-tutorials. All of the lecture-tutorials differ more or less from their fall 2009 versions, although many of these revisions involve rewordings, reordering of questions, and other minor changes meant to improve the readability and flow of each activity. I will not detail all of these minor changes here. Instead, I focus on the revisions we made in response to the data we collected in the fall of 2009.

Table 6.26 shows that the distribution of scores for item 4 on Form D did not really change pre- to post-instruction. Part of this may have to do with the wording of item 4, which is why we revised it for the spring 2010 version of Form D, as described above. However, we also added and changed several questions in the "Dark Matter" lecture-tutorial to further emphasize the relevant physical relationships. For example, after the table at the beginning of the lecture-tutorial, the fall 2009 version presents students with the following question:

> 1) Fill in the blanks to complete the following sentences. It may be helpful to base your responses on the information provided in the table above.
>
> *There are ____ planets inside Neptune's orbit and ____ planets inside Mercury's orbit. The interior mass for Neptunes orbit is ____* (much greater than/approximately the same as/much less than) *the interior mass of Mercurys orbit.*

In contrast, the spring 2010 version expands on this point by adding three additional questions:

> 1) Where is the vast majority of mass in the solar system located? What object or objects account for most of this mass?
>
> 2) How does the orbital speed of planets farther from the Sun compare to the orbital speed of planets closer to the Sun?
>
> The mass inside a planet's orbit affects how fast the planet moves because it affects the strength of the gravitational force felt by the planet. In addition, the size of the planet's orbit can affect the strength of the gravitational force exerted on the orbiting planet and therefore affects the planet's orbital speed.
>
> 3) How does the gravitational force on planets farther from the Sun compare to the gravitational force on planets closer to the Sun? Explain your reasoning.

> 4) Complete the blanks in the sentences of the following paragraph by either writing in the necessary information or circling the correct response. It may be helpful to base your responses on the information provided in the table above and your answers to the previous questions.
>
> *There are ____ planets inside Neptune's orbit and ____ planets inside Mercury's orbit. However, the interior mass for Neptune is ____* (much greater than/approximately the same as/much less than) *the interior mass of Mercury. Neptune is ____* (much closer to/much farther from/about the same distance from) *the Sun as/than Mercury. Therefore the gravitational force exerted on Neptune is ____* (stronger/weaker/about the same strength) *as/than the force exerted on Mercury. As a result, Neptune has an orbital speed that is ____* (much slower, much faster, about the same speed) *as/than the orbital speed of Mercury.*

These exemplify the kinds of revisions we made to the "Dark Matter" lecture tutorial in order to help students understand how rotation curves relate to the amount of mass inside an object's orbit.

One of the consistent findings of this chapter is that the lecture-tutorials apparently did not improve students' performances on Form A. We found this result when we looked at item P-values, normalized gains, and IRT gains. Since Form A focuses on the *Hubble plots* construct, we made several revisions to the "Hubble's Law" lecture-tutorial. First, we added the explicit statement to the lecture-tutorial after Hubble's law is introduced: "In a Hubble plot, the *expansion rate* is indicated by the slope of the graph." We then rewrote many of the questions following this statement (questions 8-16) such that they focused more on the relationship between the slope of a Hubble plot, the expansion rate of the universe, and the age of the universe. For example, after students encounter the Hubble plot for our universe (in which the expansion is accelerating; see Perlmutter *et al* 1999 and Reiss *et al* 1998), the lecture-tutorial asks the following three questions (note that Figure 5 is the Hubble plot for an accelerating universe):

> 13) Based on the straight line drawn in the Hubble plots shown in Figures 2-4, you might infer that the expansion rate for the universe is constant. Based on the Hubble plot shown in Figure 5, would you say that the expansion rate of the universe is constant or changing over time? Explain your reasoning.
>
> 14) Based on the Hubble plot in Figure 5, is the expansion rate represented by the motion of galaxies far away from us faster than, slower than, or the same as the expansion rate represented by motion of nearby galaxies? Explain your reasoning.

> 15) Based on the Hubble plot in Figure 5, is the expansion rate of the universe increasing or decreasing as time goes on? Explain your reasoning.

We anticipated that these questions and other like them would help improve the lecture-tutorial students' performances on Form A in the spring of 2010.

One issue we did not address in our revisions is related to students' performances on item 6 on Form B. This item probes whether students think the distances between planets in solar systems and stars in galaxies also grow with the expansion of the universe along with the distances between galaxies. In Section 6.3.2 we found that this item had a somewhat low discrimination and in Sections 6.3.4 and 6.5.2 we found that the lecture-tutorial students did not improve much more than the non-lecture-tutorial students on this item. The concept probed by item 6 is primarily addressed by the "Making Sense of the Universe and Expansion" lecture-tutorial. There are two reasons we did not revise this lecture-tutorial based on the data for item 6. First, this concept is covered in more detail in one of the current *Lecture-Tutorials for Introductory Astronomy* called "The Expansion of the Universe" (Prather *et al* 2008). Instructors can use this lecture-tutorial if they are concerned about this issue. Second, the way item 6 was written may have been part of the problem which is why we revised it for the spring 2010, as described above. Our data for the spring 2010 addresses whether or not the lecture-tutorial students outperform the non-lecture-tutorial students on this item post-revision.

In general, we did not make many major changes to the lecture-tutorials other than those described in this section. This is primarily due to the fact that, for Forms B-D, we have evidence that the students who used the lecture-tutorials did better than the students who did not.

6.8 Summary of Fall 2009 Results

As noted at its very beginning, this chapter focused on several key questions. First, what are Astro 101 students' common difficulties with cosmology? Second, do the lecture-tutorials help? Third, are our surveys both valid and reliable? The data we collected in the fall of 2009 helped us begin to answer all three of these questions.

What are Astro 101 students' common difficulties with cosmology? Our surveys revealed several difficulties Astro 101 students experience with cosmology. For example, we found that Hubble plots are especially hard for Astro 101 students to interpret. They frequently chose the wrong Hubble plot for a given physical situation for reasons ranging from equating a "constant rate of expansion" with a "constant velocity for galaxies" to referring to the height of a graph when they should look at its slope. Even when students select an appropriate graph, they struggle to articulate complete and correct reasons for their choices. Correctly interpreting Hubble plots is one of the most pernicious difficulties Astr 101 students experience.

We also found several issues with students' conceptualizations of the expansion and evolution of the universe. Some students, especially pre-instruction, do not even think that the universe is really expanding. Instead, they regard "the expansion of the universe" as a metaphor for how our knowledge of the universe increases over time and/or for how how new objects are created in the universe over time. We found that many students think of the Big Bang as an explosion of pre-existing matter into empty space, a belief that is often accompanied by belief in a center of the universe and an edge to the distribution of galaxies. A minority of students think that the Big Bang refers to the beginning of something smaller than the universe, such as the formation of Earth, or an event that occurred to something smaller than the universe, such as the collision of an asteroid with the Earth at the end of the Cretaceous. The effects of the expansion of the universe on its evolution are also difficult for Astro 101 students. Many students struggle to account for the effects of expansion on light travel times, lookback times, and distances, a task whose difficulty is only compounded when students think light-years are a measure of time. Many students also do not know how the temperature, density, and amount of matter in the universe change as the universe expands, although their written answers do not indicate there are any robust misconceptions that account for the responses of a majority of students. Students ideas about these topics do change pre- to post-instruction as discussed in Section 6.5 above.

Finally, we found that students have trouble even identifying the correct rotation curve for a spiral galaxy. Recognizing the fact that spiral galaxies have flat rotation curves is a necessary first step in describing while these rotation curves are evidence for dark matter. While this difficulty is not surprising for students pre-instruction (the majority of whom probably never saw a rotation curve prior to Astro 101), this difficulty appears to persist post-instruction for students who did not do the "Dark Matter" lecture-tutorial.

Do the cosmology lecture-tutorials help? In general, yes. Multiple analyses provide evidence that students who use the new suite of cosmology lecture-tutorials outperform their peers who do not on Forms B-D. For example, we see greater shifts in item P-values and the distribution in scores for the lecture-tutorial classes than the non-lecture-tutorial class. Furthermore, we also find that the lecture-tutorial students exhibit larger normalized gains on Forms B-D than the non-lecture-tutorial students. The increase of scores pre- to post-instruction is statistically significant for the lecture-tutorial students on Forms B-D, according to the Mann-Whitney test, but not for the non-lecture-tutorial students. Finally, the differences between the IRT ability gains of the lecture-tutorial and non-lecture-tutorial students are 0.34 ± 0.14 logits for Form B and 0.70 ± 0.14 logits for Form C. Taken together, these data are consistent with the hypothesis that the lecture-tutorials improve students' knowledge of the constructs covered by Forms B-D.

Of course, the story is different for Form A. None of the above approaches for looking at our data detected any evidence that the lecture-tutorial students performed better on Form A than the non-lecture-tutorial students. This null result inspired many of our revisions described in Section 6.7.

Are the conceptual cosmology surveys valid and reliable? We have evidence for both. In Sections 6.3.3 and 6.4.5 we look at the reliability of our surveys from both classical test theory and item response theory perspectives. Our surveys' values of Cronbach's α are smaller than we prefer; however, this may be due to the brevity of our surveys. Inter-rater reliability and IRT reliability analyses both support the reliability of our surveys. Finally, Section 6.6 presents evidence for our interpretive argument for the validity of our surveys. Although this evidence does not prove the validity of our surveys in any absolute sense, it is part of a larger argument continued in Chapters 7 and 8 meant to support our assertion that we created valid surveys for this study.

Finally, we described the revisions we made to both our surveys and our lecture-tutorials as a result of the data presented in this chapter. We gave these revised surveys and lecture-tutorials to Astro 101 students in the spring 2010 semester. The following chapter presents our analysis of the data we collected in the spring of 2010.

Chapter 7

Spring 2010 Results

We continued our study in the spring 2010 semester with students drawn from four different Astro 101 classes taught at three separate institutions. We administered the revised versions of our surveys (see Appendix D for copies of the surveys and Appendix F for the associated scoring rubrics) to all of these classes pre- and post-instruction. Two of these classes also used the revised versions of the lecture-tutorials described in Section 6.7.2. This chapter presents the results from this part of our study.

In many cases, the spring 2010 results are similar to the fall 2009 results. In others, the spring 2010 data add new insights to our understanding of students' difficulties with cosmology and the efficacy of the cosmology lecture-tutorials. This is especially true for items we revised after the fall 2009 semester and for items that address topics covered by the revised sections of the lecture-tutorials. I will highlight the similarities and differences between the fall 2009 and spring 2010 data when they are significant and relevant. However, much of the analysis described below assumes the reader is familiar with the techniques, processes, and results described in previous chapters, especially Chapter 6.

This chapter follows the same basic outline as Chapter 6. Section 7.1 briefly describes each of the four participating Astro 101 classes. Section 7.2 contains our CTT analysis of our data. Section 7.3 contains our IRT analysis. Section 7.4 examines students' responses to our survey items in detail. Section 7.5 continues and expands our validity argument begun in the previous chapter. Section 7.6 then describes the set of revisions we made to the surveys and the lecture-tutorials prior

Table 7.1: Number of participants pre- and post-instruction per class for spring 2010.

Class	Pre-Instruction	Post-Instruction
Class D	687	626
Class E	237	220
Class F	136	86
Class G	155	149

to our final semester of data collection (fall 2010), which is the subject of Chapter 8. Section 7.7 summarizes the findings of this chapter.

7.1 Surveyed Classes

As noted above, four Astro 101 classes participated in our study during the spring 2010 semester. Class D was taught at the University of Arizona. It was identical in all major respects (including the instructor and use of lecture-tutorials) to Class A from the fall 2009, except Class D was larger by several hundred students. Classes E and F were both taught at the University of Colorado at Boulder. Class E did not have any laboratories or recitation sections, but its students did use the lecture-tutorials. Class F, like Class B from the fall 2009, had weekly mandatory small group recitation sections run by undergraduate learning assistants. Unlike Class B, Class F did not use any lecture-tutorials. The instructor of Class F relied primarily on lecture to teach cosmology. Finally, Class G was taught at Syracuse University. Although the instructor of Class G was familiar with the original *Lecture-Tutorials for Introductory Astronomy* (Prather *et al.* 2008), he did not use the cosmology lecture-tutorials in his course. Table 7.1 shows the number of participating students from each class, pre- and post-instruction.

As in the fall 2009, all four surveys were administered at once pre- and post-instruction, meaning only about a quarter of each class took a particular form during a given administration. Also like the fall 2009, we will not look at any matched pre- and post-instruction data.

Table 7.2: The discriminations of the items on Forms A-D for spring 2010.

Form A		Form B		Form C		Form D	
Item	Discrimin.	Item	Discrimin.	Item	Discrimin.	Item	Discrimin.
Item 1	0.63	Item 1	0.57	Item 1	0.64	Item 1	0.58
Item 2	0.65	Item 2	0.57	Item 2	0.54	Item 2	0.57
Item 3	0.54	Item 3	0.46	Item 3	0.62	Item 4	0.84
Item 4	0.53	Item 4	0.54	Item 4	0.54		
Item 5	0.60	Item 5	0.61	Item 5	0.52		
Item 6	0.63	Item 6	0.55	Item 6	0.59		
		Item 7	0.47				

7.2 Classical Test Theory Analysis

This section presents our CTT analysis of the spring 2010 data. Like Section 6.3 from the previous chapter, we use CTT to examine the difficulties and discriminations of our survey items (Section 7.2.1), evaluate the reliabilities of our surveys (Section 7.2.2), and calculate normalized gains for each class for each survey form (Section 7.2.3). Readers who require an overview of classical test theory should turn back to Section 6.3.1.

7.2.1 Items' Difficulties and Discriminations

Table 7.2 shows the discriminations (i.e. the correlation between item scores and total test scores) for each item on Forms A-D for the spring 2010. All item discrimination values lie within conventionally accepted limits and most are relatively high (> 0.50), the exceptions being items 3 and 7 from Form B. Note, however, that the discrimination of item 7 is larger than the discrimination of its fall 2009 version (when it was item 6 on Form B; see Table 6.2), suggesting that our revisions to this item helped.

Table 7.3 shows the P-values for the items of Forms A-D for the spring 2010 semester. For each item, we show its overall P-value, and its P-values using just the lecture-tutorial students pre-instruction ("LT Pre"), the lecture-tutorial students post-instruction ("LT post"), the non-lecture-tutorial students pre-instruction ("Non-LT Pre"), and the non-lecture-tutorial students post-instruction ("Non-LT Post"), just as we did in Chapter 6. All of these P-values fall within

Table 7.3: The difficulties (P-values) of the items on Forms A-D for spring 2010.

	Item	Overall	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Form A	Item 1	0.50	0.49	0.53	0.43	0.52
	Item 2	0.42	0.41	0.45	0.35	0.42
	Item 3	0.36	0.35	0.39	0.32	0.32
	Item 4	0.35	0.34	0.38	0.32	0.33
	Item 5	0.82	0.78	0.89	0.73	0.84
	Item 6	0.40	0.34	0.50	0.31	0.39
Form B	Item 1	0.50	0.37	0.63	0.46	0.69
	Item 2	0.66	0.53	0.84	0.57	0.62
	Item 3	0.81	0.80	0.82	0.78	0.81
	Item 4	0.42	0.37	0.52	0.36	0.40
	Item 5	0.50	0.44	0.58	0.45	0.51
	Item 6	0.61	0.61	0.68	0.45	0.56
	Item 7	0.69	0.64	0.74	0.61	0.81
Form C	Item 1	0.64	0.44	0.90	0.47	0.72
	Item 2	0.54	0.43	0.72	0.41	0.49
	Item 3	0.62	0.49	0.81	0.47	0.64
	Item 4	0.54	0.48	0.66	0.43	0.50
	Item 5	0.52	0.49	0.57	0.48	0.51
	Item 6	0.59	0.60	0.56	0.60	0.62
Form D	Item 1	0.64	0.52	0.79	0.54	0.64
	Item 2	0.87	0.85	0.89	0.91	0.86
	Item 4	0.59	0.54	0.65	0.50	0.64

conventionally accepted limits. Furthermore, by comparing the shifts between the pre- and post-instruction P-values for the lecture-tutorial and non-lecture-tutorial students on each item, we find many items for which the lecture-tutorial students exhibit larger changes than the non-lecture tutorial students (items 3, 4, and 6 on Form A, items 1, 2, 4, and 5 on Form B, items 1-5 on Form C, and items 1 and 2 on Form D). However, there are also items for which the non-lecture-tutorial population shows a greater change than the lecture-tutorial population (items 1 and 2 on Form A, items 3, 6, and 7 on Form B, item 6 on Form C, and item 4 on Form D). Finally, the change in P-values pre- to post-instruction is identical for both populations for item 5 on Form A. We examine these patterns in more detail and from different perspectives in subsequent sections and attempt to deduce which changes are significant and which are not.

Table 7.4: Cronbach's α for Forms A-D for spring 2010.

Form	Cronbach's α
Form A	0.62
Form B	0.57
Form C	0.56
Form D	0.27

7.2.2 Reliability

Table 7.4 shows Cronbach's α for each form for the spring 2010 semester. A cursory comparison with the analogous table in Chapter 6 (Table 6.4) shows the value of Cronbach's α increased for Form A, decreased only slightly for Form B, and decreased significantly for Forms C and D. There are two competing effects that may explain these changes. First, the fact that we added additional items to Forms A-C should have raised their values of Cronbach's α since Cronbach's α is sensitive to the length of the test. However, Form A was the only survey whose value of Cronbach's α increased from the fall 2009 to the spring 2010. What about the other surveys? As discussed in Chapter 6, Cronbach's α also depends on the homogeneity of the test taker population. More homogeneous populations necessarily produce smaller values of Cronbach's α. This may explain why we see a small decrease in Form B's Cronbach's α and larger decreases in the values of Cronbach's α for Forms C and D. In some cases, the item P-values show that the items we added were of equal difficulty for both populations of students, pre- and post-instruction. Furthermore, the fact that we have a larger sample of non-lecture-tutorial students may also depress the values of Cronbach's α for these surveys since we might expect these students to exhibit smaller gains on many items than their peers who used the lecture-tutorials. These homogenizing influences may account for the lower values of Cronbach's α this semester compared to the fall 2009. Whatever the cause, Table 7.4 underscores the sample-dependent nature of Cronbach's α (Schmitt 1996; Thompson 2003) and supports our decision to look at other measures of reliability.

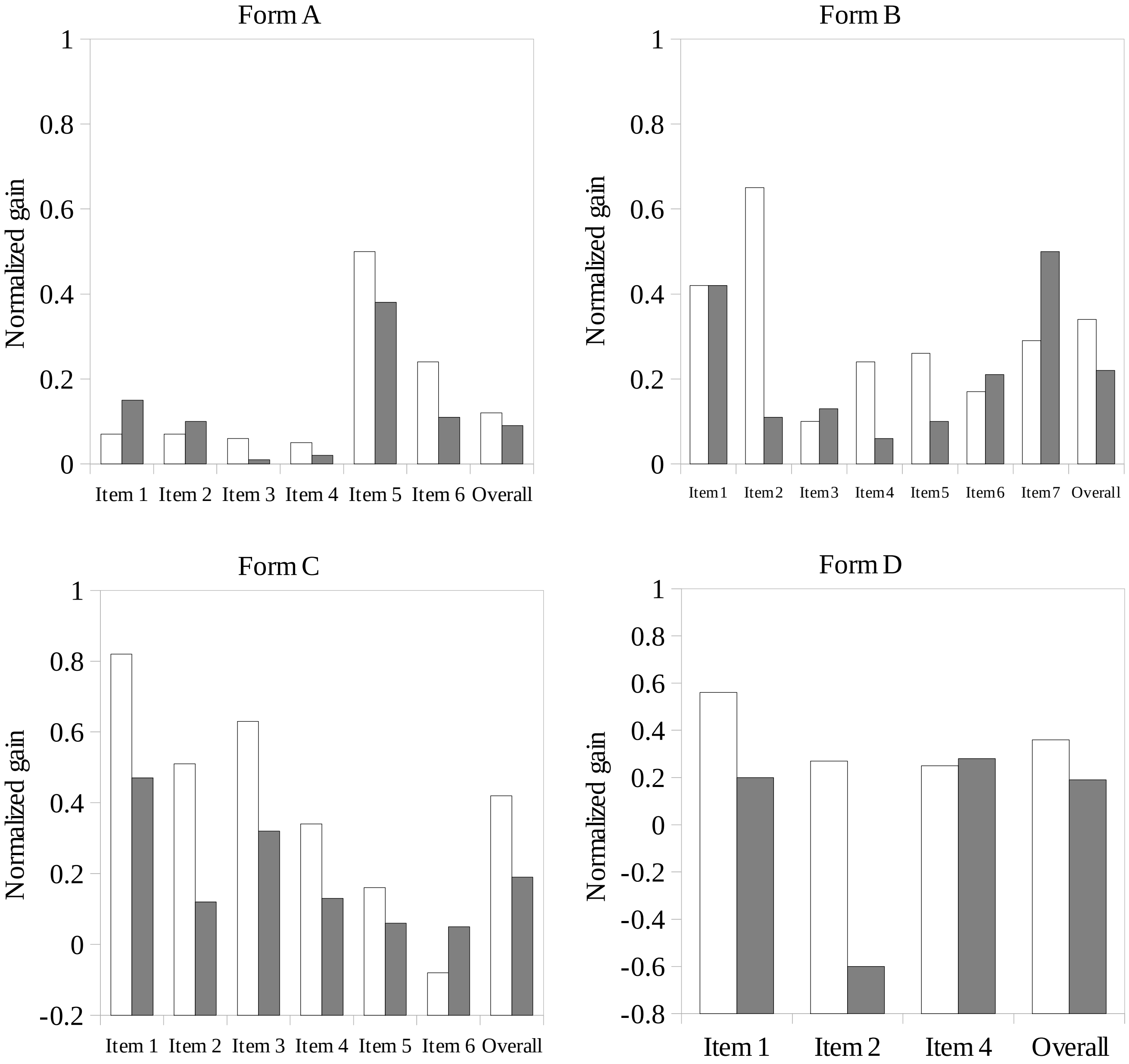


Figure 7.1: The normalized gains for Forms A-D for spring 2010. White bars correspond to the lecture-tutorial population and grey bars correspond to the non-lecture-tutorial population.

7.2.3 Normalized Gains

Figure 7.1 shows the normalized gains for each item on each survey, as well as for Forms A-D overall, for both the lecture-tutorial and non-lecture-tutorial populations. In general, the lecture-tutorial students have larger gains than their non-lecture-tutorial counterparts, with a handful of

Table 7.5: Mann-Whitney p-values for the lecture-tutorial (LT) and non-lecture-tutorial (Non-LT) groups for Forms A-D for spring 2010.

Population	Form A	Form B	Form C	Form D
LT	< 0.0001	< 0.0001	< 0.0001	< 0.0001
Non-LT	0.0017	< 0.0001	< 0.0001	0.0275

items (items 1 and 2 on Form A, items 1, 3, 6, and 7 on Form B, item 6 on Form C, and item 4 on Form D) as exceptions. Furthermore, the non-lecture-tutorial students' gains are, in many cases, very low; this is true for several of the items for which they exhibit larger gains than the lecture-tutorial students. An examination of the overall gains for each form shows that the lecture-tutorial students always have the larger gains, although the significance of this results is not immediately obvious in some cases (for example, Form A).

Table 7.5 gives the Mann-Whitney p-values for each form for both the lecture-tutorial and non-lecture-tutorial populations. These values show that the gains in Figure 7.1 are statistically significant ($p < 0.05$) for both populations on all four forms. In Section 7.3.5 below, we use IRT-estimated abilities to evaluate whether or not the differences in gains are statistically significant for Forms A-C.

7.3 Item Response Theory Analysis

In this section, we apply the partial credit IRT model (Masters 1982) to our data from Forms A-C. This section follows the same outline as Section 6.4 in Chapter 6: Section 7.3.1 discusses the item's step difficulties and Thurstonian thresholds, Section 7.3.2 looks at how well our data meets the assumptions of item response theory, Section 7.3.3 examines how well the partial credit model fits our data, Section 7.3.4 addresses the reliability of Forms A-C from an IRT perspective, and Section 7.3.5 presents the gains for the lecture-tutorial and non-lecture-tutorial students calculated from their IRT-estimated abilities. See Section 6.4.1 for an overview of IRT in general and the partial credit model in particular.

7.3.1 Items' Difficulties

Table 7.6: The step difficulty b_{ij} and Thurstonian Threshold β_j parameters for the items on Forms A-C for spring 2010. All values are in logits.

		Step Parameters				Thurstonian Thresholds			
	Item	b_{01}	b_{12}	b_{23}	b_{34}	β_1	β_2	β_3	β_4
Form A	Item 1	-5.36	0.17	3.84	-	-5.37	0.15	3.86	-
	Item 2	-4.75	1.36	5.37	-	-4.76	1.34	5.38	-
	Item 3	-5.21	3.42	3.11	-	-5.21	2.85	3.68	-
	Item 4	-4.95	3.40	3.46	-	-4.95	2.94	3.93	-
	Item 5	-3.62	-0.95	-	-	-3.68	-0.88	-	-
	Item 6	-2.55	1.20	4.36	-	-2.58	1.18	4.40	-
Form B	Item 1	-1.60	-0.27	3.38	-0.17	-1.80	-0.12	1.60	1.75
	Item 2	-1.24	-0.66	1.72	-	-1.58	-0.41	1.80	-
	Item 3	-1.60	-0.19	-	-	-1.78	0.00	-	-
	Item 4	-2.85	1.19	2.85	-	-2.86	1.66	3.12	-
	Item 5	-2.04	2.20	0.52	-	-2.05	1.17	1.59	-
	Item 6	-0.40	1.31	-0.25	-	-0.62	0.51	0.86	-
	Item 7	-1.78	0.83	-	-	-1.84	0.90	-	-
Form C	Item 1	-2.67	1.89	-0.11	-	-2.68	0.73	1.09	-
	Item 2	-2.06	1.32	1.45	-	-2.09	0.91	1.90	-
	Item 3	-1.58	2.07	-0.30	-	-1.61	0.79	1.07	-
	Item 4	-3.49	2.83	-0.17	-	-3.48	1.22	1.45	-
	Item 5	-2.77	2.31	0.56	-	-2.77	1.23	1.64	-
	Item 6	-1.32	1.69	0.33	-	-1.38	0.83	1.29	-

The step difficulties and Thurstonian thresholds for each item on the spring 2010 versions of Forms A-C are given in Table 7.6. As with the fall 2009 items, many of the step difficulties and Thurstonian thresholds for the items shown in Table 7.6 fall between the -3 and 3 logits. We are not concerned with values lower than -3 since we constructed our scoring rubrics (see Appendix F) such that the requirements for earning a score of 1 instead of 0 on many items are minimal. However, the high values of b_{23} and β_3 for items 1-4 and 6 on Form A once again grab our attention. Even though these values are smaller than their fall 2009 counterparts (see Table 6.6), they still raise questions about whether students are interpreting these items as we intended and/or whether they are too difficult for many students (even lecture-tutorial students post-instruction). Section 7.4.1, in which we look at what students are actually saying in their responses, and Section 7.5, in which we discuss the responses of interviewed students as part of our validity argument, will help answer these questions.

7.3.2 Testing Item Response Theory's Assumptions

In this section we look at whether or not our data meets the two assumptions of item response theory – unidimensionality and parameter invariance – and try to make sense out of any deviations we detect. Following our approach in the previous chapter, we use the item outfit statistics (see the following section) to detect potential departures from unidimensionality (Smith and Miao 1994). Since all but one of the item outfit statistics fall within their theoretically accepted ranges for Forms A and B (see Section 7.3.3 below), we have no evidence for any multidimensional structure for these forms. Form C is somewhat more complicated. As discussed in the following section, all but one of Form C's items possess outfit values outside of the expected range. This result encourages us to exercise caution in claiming all the advantages of IRT for Form C.

Our investigation of parameter invariance also gives us reason to be cautious with Form C. We again used Yen's Q3 statistic (Yen 1984) to detect potential violations of parameter invariance, flagging all item pairs for which the Q3 statistic exceeds |0.20|. Tables 7.7-7.9 show Yen's Q3 statistic for each item pair on Forms A-C, respectively. There are two item pairs on Form A whose Q3 statistic is greater than |0.20|. Both item pairs make sense. Item 1 asks about a universe expanding at a constant rate while item 2 asks about a universe contracting at a constant rate. Item 3 asks about a universe expanding at a faster and faster rate over time while item 4 asks about a universe expanding at a slower and slower rate over time. In both cases, many students simply give the "opposite" reason for one item that they gave for the other (see Section 7.4.1). We did not flag any item pairs on Form B. On Form C, we flagged four different pairs of items. This suggests that Form C does not obey the assumption of parameter invariance.

7.3.3 Model Fit

Table 7.10 shows the outfit statistics for the items on Forms A-C. Following Wu and Adams (2011), we expect Form A's items to have outfit values between 0.88 and 1.12, Form B's items to

Table 7.7: Yen's Q3 statistic for each pair of items on Form A for the spring 2010.

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Item 1	1	**0.41**	0.09	0.09	0.02	0.03
Item 2	**0.41**	1	0.18	0.14	-0.07	0.11
Item 3	0.09	0.18	1	**0.77**	0.01	0.08
Item 4	0.09	0.14	**0.77**	1	0.08	0.08
Item 5	0.02	-0.07	0.01	0.08	1	-0.09
Item 6	0.03	0.11	0.08	0.08	-0.09	1

Table 7.8: Yen's Q3 statistic for each pair of items on Form B for the spring 2010.

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Item 1	1	-0.02	0.01	0.03	0.04	-0.01	0.19
Item 2	-0.02	1	0.07	0.09	-0.03	0.02	0.05
Item 3	0.01	0.07	1	-0.13	-0.15	0.00	0.06
Item 4	0.03	0.09	-0.13	1	0.16	-0.04	0.15
Item 5	0.04	-0.03	-0.15	0.16	1	0.11	0.08
Item 6	-0.01	0.02	0.00	-0.04	0.11	1	-0.04
Item 7	0.19	0.05	0.06	0.15	0.08	-0.04	1

Table 7.9: Yen's Q3 statistic for each pair of items on Form C for the spring 2010.

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Item 1	1	0.19	**0.34**	**0.22**	0.13	0.03
Item 2	0.19	1	0.16	0.15	0.12	-0.02
Item 3	**0.34**	0.16	1	**0.28**	0.11	0.00
Item 4	**0.22**	0.15	**0.28**	1	**0.22**	0.03
Item 5	0.13	0.12	0.11	**0.22**	1	-0.03
Item 6	0.03	-0.02	0.00	0.03	-0.03	1

have outfit values between 0.89 and 1.12, and Form C's items to have outfit values between 0.89 and 1.11. On Form A, only item 1 falls outside the range of expected values. Form B has no items with outfit values outside the range of 0.89 to 1.12. However, all of Form C's items, save item 5, fall outside of its expected range. What does this mean? In all cases, the departure from the expected range is slight. Furthermore, the fact that in all but one instance the outfit values lie above the maximum expected outfit implies that the variances in the observed residuals are greater than expected – which means that the items are *easier* than the partial credit model predicts (Wilson 2005).

Table 7.10: Outfit statistics for the items on Forms A-C for the spring 2010.

Item	Form A	Form B	Form C
Item 1	**1.16**	1.00	**1.15**
Item 2	1.12	0.95	**1.19**
Item 3	1.11	1.01	**1.13**
Item 4	1.09	0.89	**1.12**
Item 5	0.90	0.97	1.10
Item 6	0.93	1.11	**0.88**
Item 7	-	1.01	-

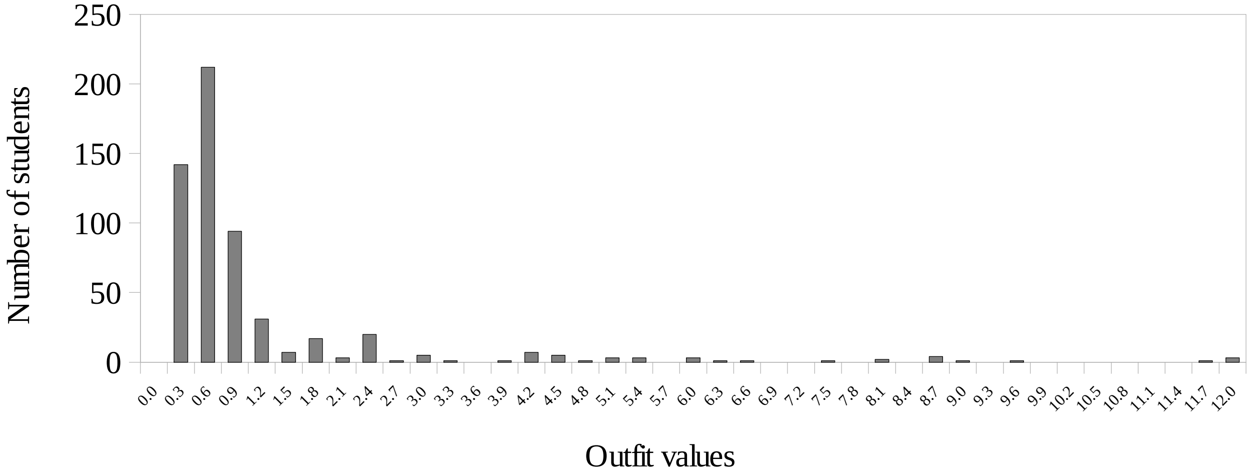


Figure 7.2: A histogram of students' outfit values for Form A for the spring 2010.

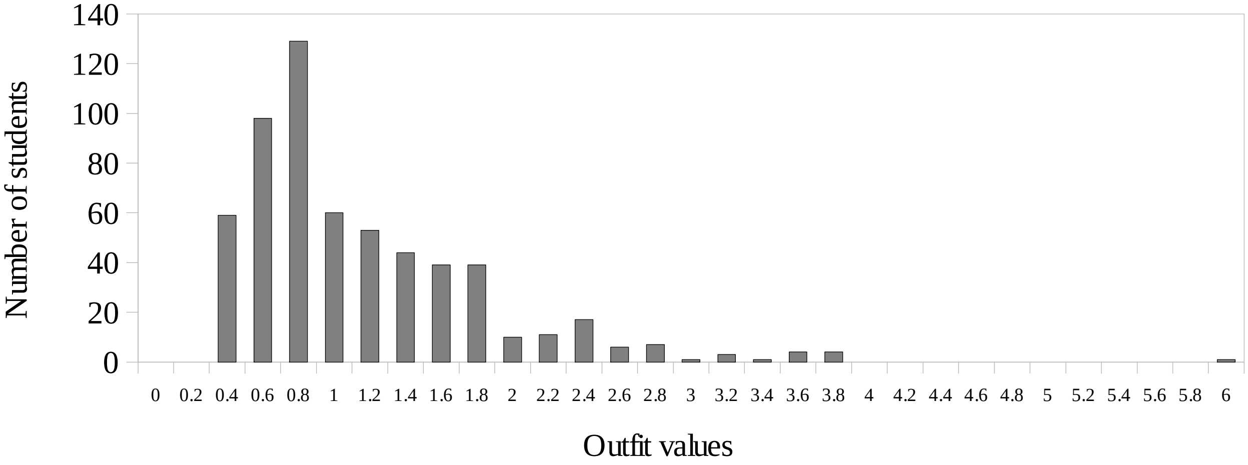


Figure 7.3: A histogram of students' outfit values for Form B for the spring 2010.

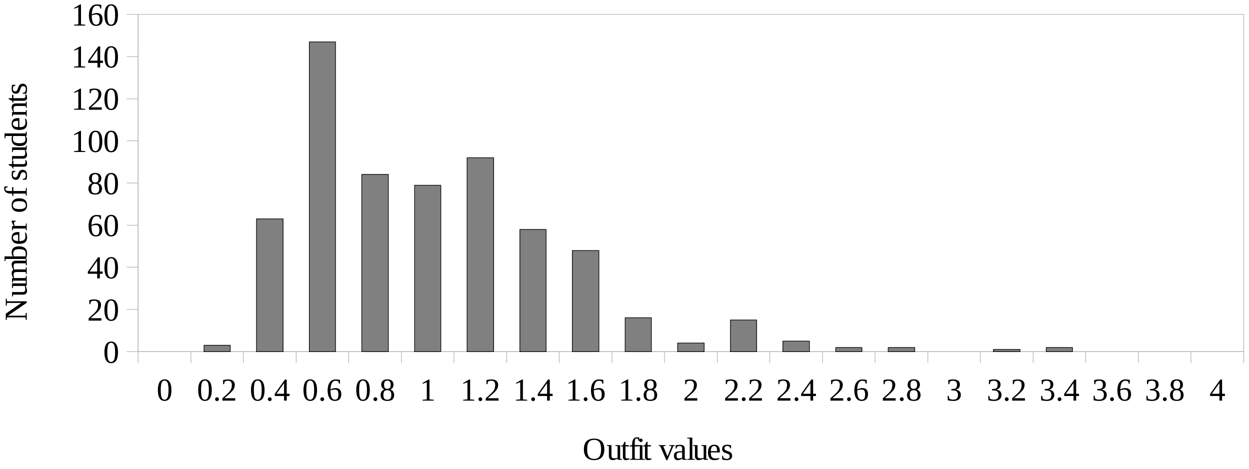


Figure 7.4: A histogram of students' outfit values for Form C for the spring 2010.

What about the outfit values for the respondents? Histograms of the respondent outfit values are shown in Figures 7.2-7.4 for Forms A-C, respectively. According to Wu and Adams (2011), we expect 95% of respondents to have outfit values less than 2.13 for Forms A and C and less than 2.05 for Form B. 87% of students fall in this range for Form A, 92% for Form B, and 97% for Form C. These numbers are almost identical to those of students in the fall 2009 semester (see Section 7.3.3) and implies that the model fit for respondents is, once again, most problematic for Form A.

7.3.4 Reliability

Can Forms A-C provide reliable estimates of students' abilities? Figure 7.5 shows how the standard error of measurement in student ability changes as a function of ability for Forms A-C. Combining this plot with the histograms of students' abilities in the Wright maps for Forms A-C (Figures 7.6-7.8) shows that, for all three forms, the minima in the standard errors roughly correspond to the peaks in the ability distributions. This, plus the fact that the Thurstonian thresholds of the items "span the space" of students' abilities on all three Forms (see Figures 7.6-7.8), supports the notion that Forms A-C are reliable.

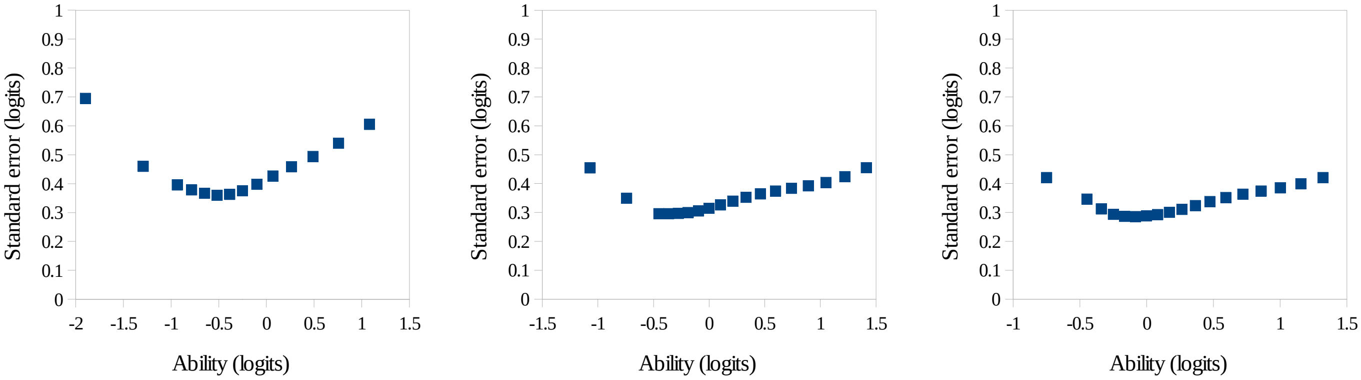


Figure 7.5: Standard error of measurement as a function of ability for (from left to right) Form A, Form B, and Form C for spring 2010.

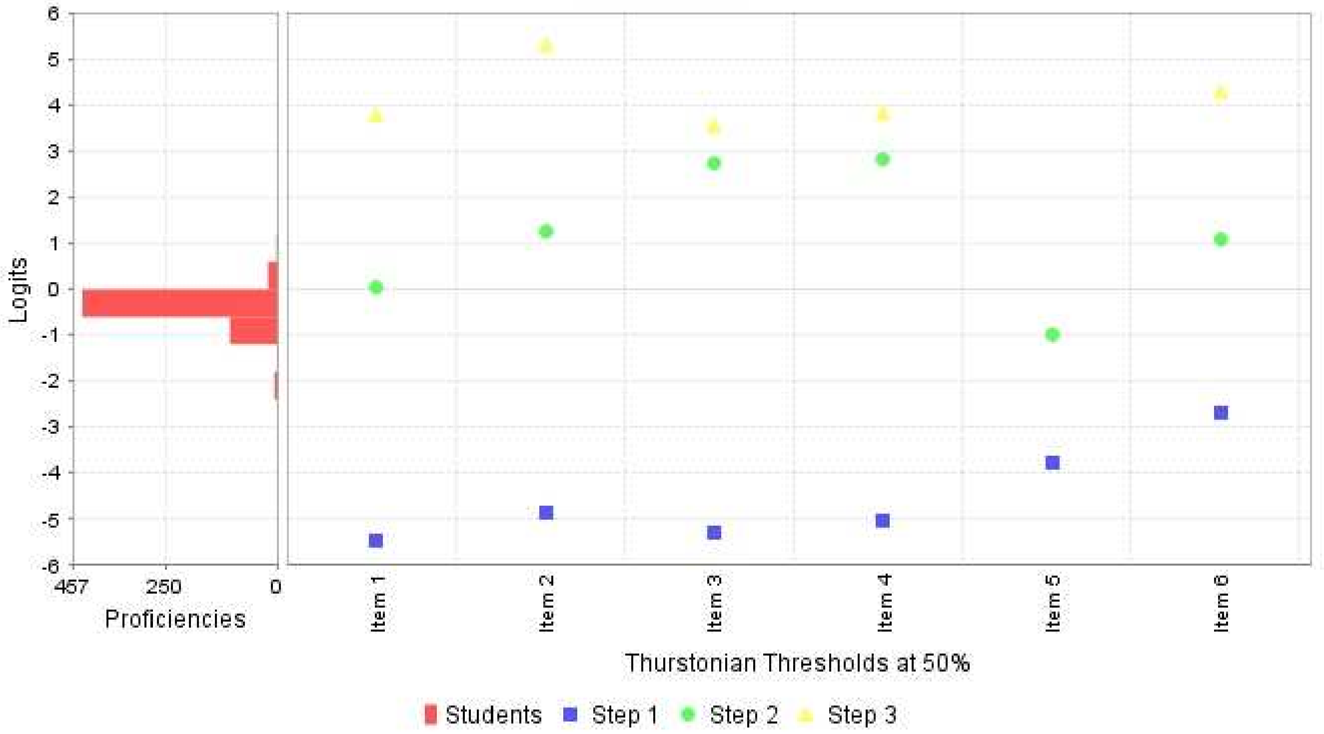


Figure 7.6: The Wright map for Form A for spring 2010.

7.3.5 IRT Gains

The fact that we detect departures from unidimensionality and parameter invariance and issues with model fit for Form C (and, to a lesser extent Form A) indicates that we cannot leverage the full theoretical advantages (such as parameter invariance) and analysis techniques offered

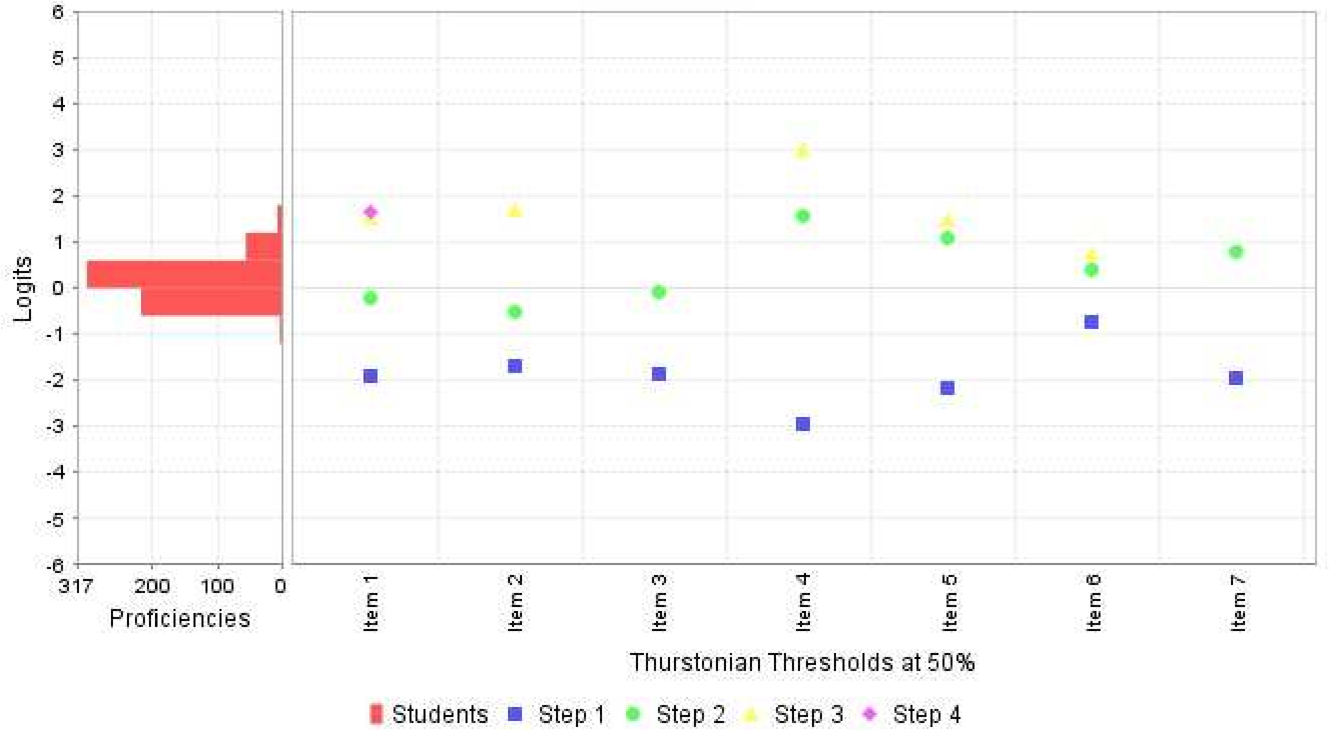


Figure 7.7: The Wright map for Form B for spring 2010.

by IRT. However, for the purposes of this study, we do not need all these advantages and techniques. We simply need to be able to make ordinal comparisons between the lecture-tutorial and non-lecture-tutorial students to see if there is any evidence that the lecture-tutorials affect student performance on Forms A-C. The fact that these forms appear to be reliable, from an IRT perspective, suggests this can be done.

Table 7.11 shows the average pre- and post-instruction abilities for both the lecture-tutorial and non-lecture-tutorial students. It also shows the gains (i.e. the difference between the average post- and pre-instruction abilities) for both populations and the difference between their gains. This last quantity is the number of interest. For Form A, we find this difference in gains is 0.04 ± 0.07 – or, in other words, we cannot conclude that using the lecture-tutorials helped students' performances on Form A compared to what they would have achieved *sans* lecture-tutorials. The story is different for Forms B and C. As Table 7.11, the difference in gains is about twice as large as its associated

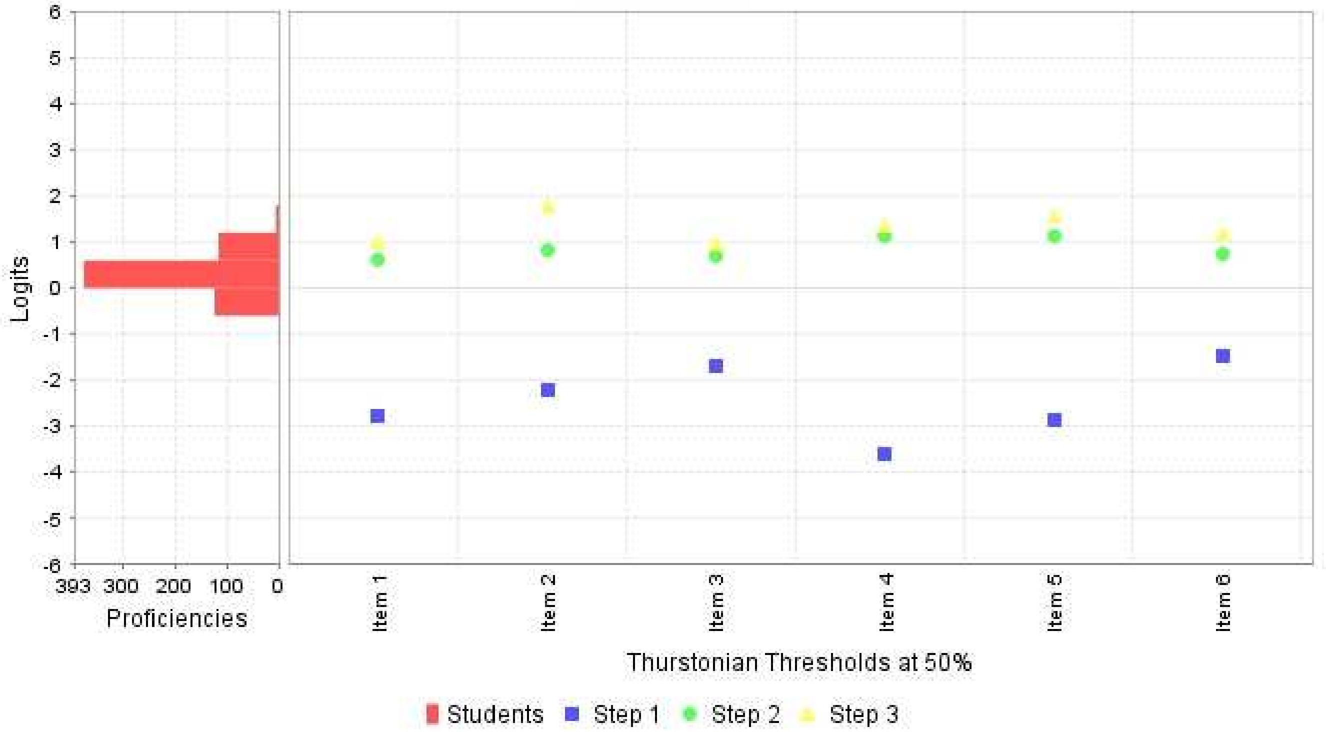


Figure 7.8: The Wright map for Form C for spring 2010.

uncertainty for Form B. For Form C, the difference in gains is about 3.7 times larger than its associated uncertainty. This suggests that the students who use the lecture-tutorials do outperform their peers who do not.

Table 7.11: Average pre-instruction IRT scores, post-instruction IRT scores, and IRT gains for the lecture-tutorial and non-lecture-tutorial classes, as well as the difference between their gains, for Forms A-C in the spring 2010. All values are in logits.

Form	LT pre	Non-LT pre	LT post	Non-LT post	LT Gain	Non-LT Gain	LT Gain-Non-LT Gain
Form A	-0.47 ± 0.02	-0.53 ± 0.04	-0.33 ± 0.03	-0.43 ± 0.05	0.14 ± 0.04	0.10 ± 0.06	0.04 ± 0.07
Form B	0.01 ± 0.02	0.02 ± 0.04	0.36 ± 0.02	0.22 ± 0.04	0.36 ± 0.03	0.20 ± 0.06	0.15 ± 0.07
Form C	0.14 ± 0.02	0.12 ± 0.03	0.53 ± 0.02	0.29 ± 0.04	0.39 ± 0.03	0.17 ± 0.05	0.22 ± 0.06

7.4 Breakdown of Item Responses

This section contains a breakdown of the responses we received from the participating students to each item, pre- and post-instruction. As such, it is meant to complement the quantitative analyses of Sections 7.2 and 7.3 with information that actually describes what students are saying in their responses to our survey items. This section is divided up into four subsections, each of which focuses on the items on one of our four surveys.

Due to the wealth of data from this semester, I employ a strategic approach to its presentation. First, although this section contains data for each individual class participating in this study during the spring 2010, I will continue the approach of the previous chapter and focus the discussion on the aggregate results for the lecture-tutorial and non-lecture-tutorial populations. This facilitates the comparison between the two groups, which, in turn, helps us address whether or not the lecture-tutorials affect how students respond to the survey items. Because there is so much response data to consider, I will also focus this discussion on points that diverge from or add to the results of the analogous section in Chapter 6. This approach should help streamline the presentation of this material while still providing interested readers with the ability to examine in detail the plethora of data we collected.

7.4.1 Form A Responses

Table 7.12 shows the distribution of scores on Form A pre-instruction and Table 7.13 shows the distribution of scores post-instruction. We scored all of Form A's items on a scale of 0-3, except for item 5 which we scored on a scale of 0-2 (see Appendix F). Tables 7.12 and 7.13 show that for the first four items, most responses fall into score categories 1 and 2. This is true for both populations, pre- and post-instruction. This implies that most student responses for items 1-4 are either incorrect or incomplete. Any gains observed on these items (see Figure 7.1) are mainly due to students moving from a score of 1 to a score of 2. The gains in students' scores on these items are small and, overall, instruction appears to have a minimal effect on students' performances on these items, regardless of whether or not that instruction included the cosmology lecture-tutorials. This is similar to the results from the fall 2009.

Table 7.12: Pre-instruction distribution of scores on Form A for spring 2010.

	Class D				Class E				LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	0%	55%	41%	4%	0%	56%	42%	2%	0%	55%	41%	3%
Item 2	1%	75%	24%	0%	2%	76%	23%	0%	1%	75%	24%	0%
Item 3	0%	94%	6%	0%	0%	97%	3%	0%	0%	95%	5%	0%
Item 4	0%	95%	5%	0%	0%	100%	0%	0%	0%	97%	3%	0%
Item 5	4%	46%	51%	-	0%	21%	79%	-	3%	39%	59%	-
Item 6	12%	75%	12%	1%	8%	79%	13%	0%	11%	76%	12%	0%
	Class F				Class G				Non-LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	0%	67%	33%	0%	11%	54%	33%	2%	6%	59%	33%	1%
Item 2	0%	85%	15%	0%	13%	74%	13%	0%	8%	78%	14%	0%
Item 3	0%	100%	0%	0%	11%	87%	2%	0%	6%	92%	1%	0%
Item 4	3%	97%	0%	0%	11%	83%	7%	0%	8%	89%	4%	0%
Item 5	0%	27%	73%	-	17%	37%	46%	-	10%	33%	57%	-
Item 6	6%	70%	24%	0%	37%	52%	11%	0%	24%	59%	16%	0%

Table 7.13: Post-instruction distribution of scores on Form A for spring 2010.

	Class D				Class E				LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	0%	43%	52%	6%	0%	52%	48%	0%	0%	45%	51%	4%
Item 2	0%	62%	36%	2%	2%	75%	23%	0%	0%	66%	32%	1%
Item 3	0%	84%	11%	5%	2%	92%	5%	2%	0%	87%	9%	4%
Item 4	1%	84%	12%	4%	3%	91%	5%	2%	1%	86%	10%	3%
Item 5	1%	18%	80%	-	3%	17%	80%	-	2%	18%	80%	-
Item 6	5%	42%	51%	2%	5%	48%	41%	6%	5%	44%	48%	3%
	Class F				Class G				Non-LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	0%	52%	43%	5%	3%	38%	59%	0%	2%	43%	53%	2%
Item 2	0%	90%	10%	0%	3%	59%	38%	0%	2%	71%	28%	0%
Item 3	0%	100%	0%	0%	5%	95%	0%	0%	3%	97%	0%	0%
Item 4	0%	100%	0%	0%	3%	95%	3%	0%	2%	97%	2%	0%
Item 5	0%	24%	76%	0%	3%	32%	65%	0%	2%	29%	69%	0%
Item 6	10%	81%	10%	0%	8%	59%	32%	0%	9%	67%	24%	0%

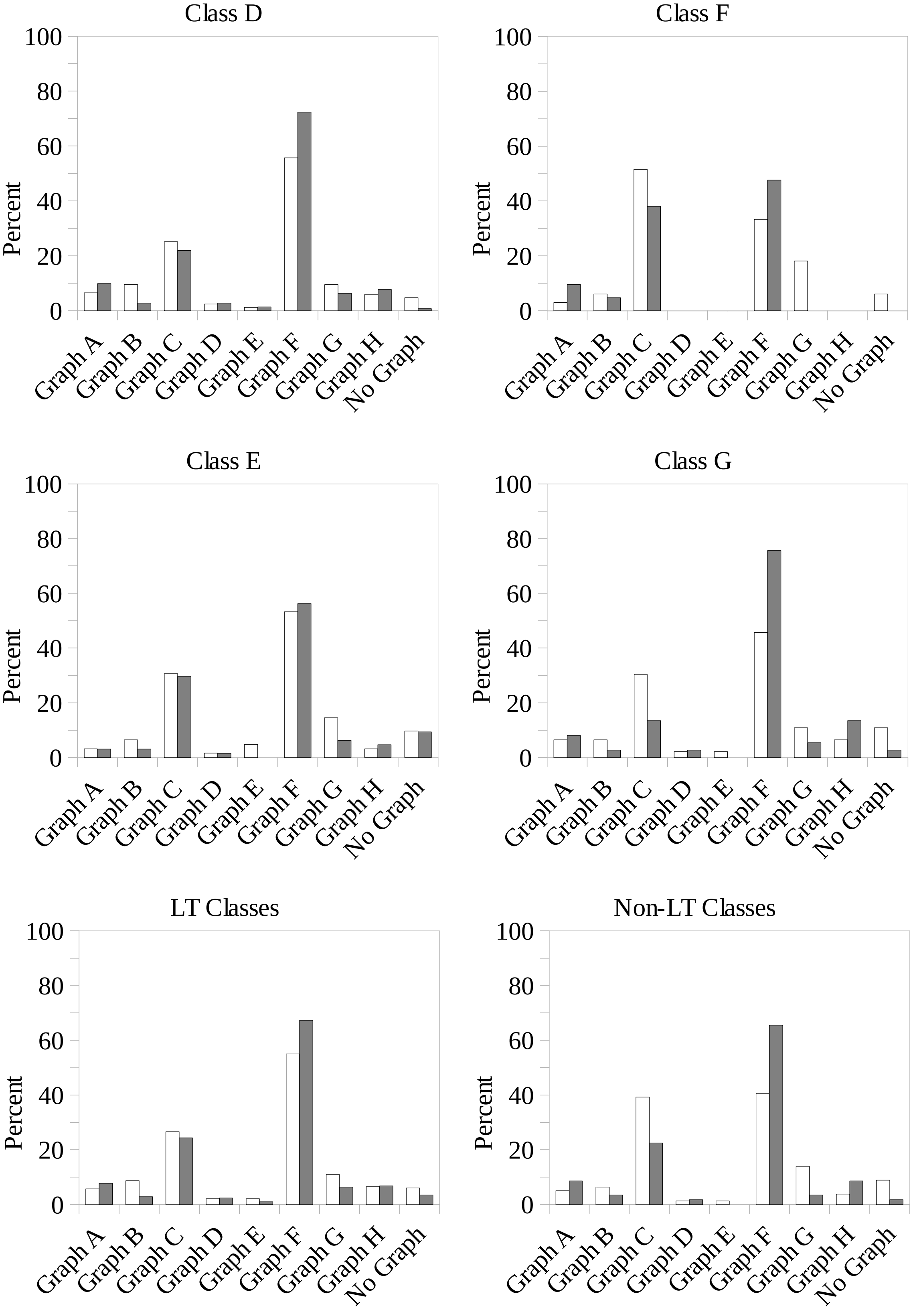


Figure 7.9: Students' graph choices for item 1 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

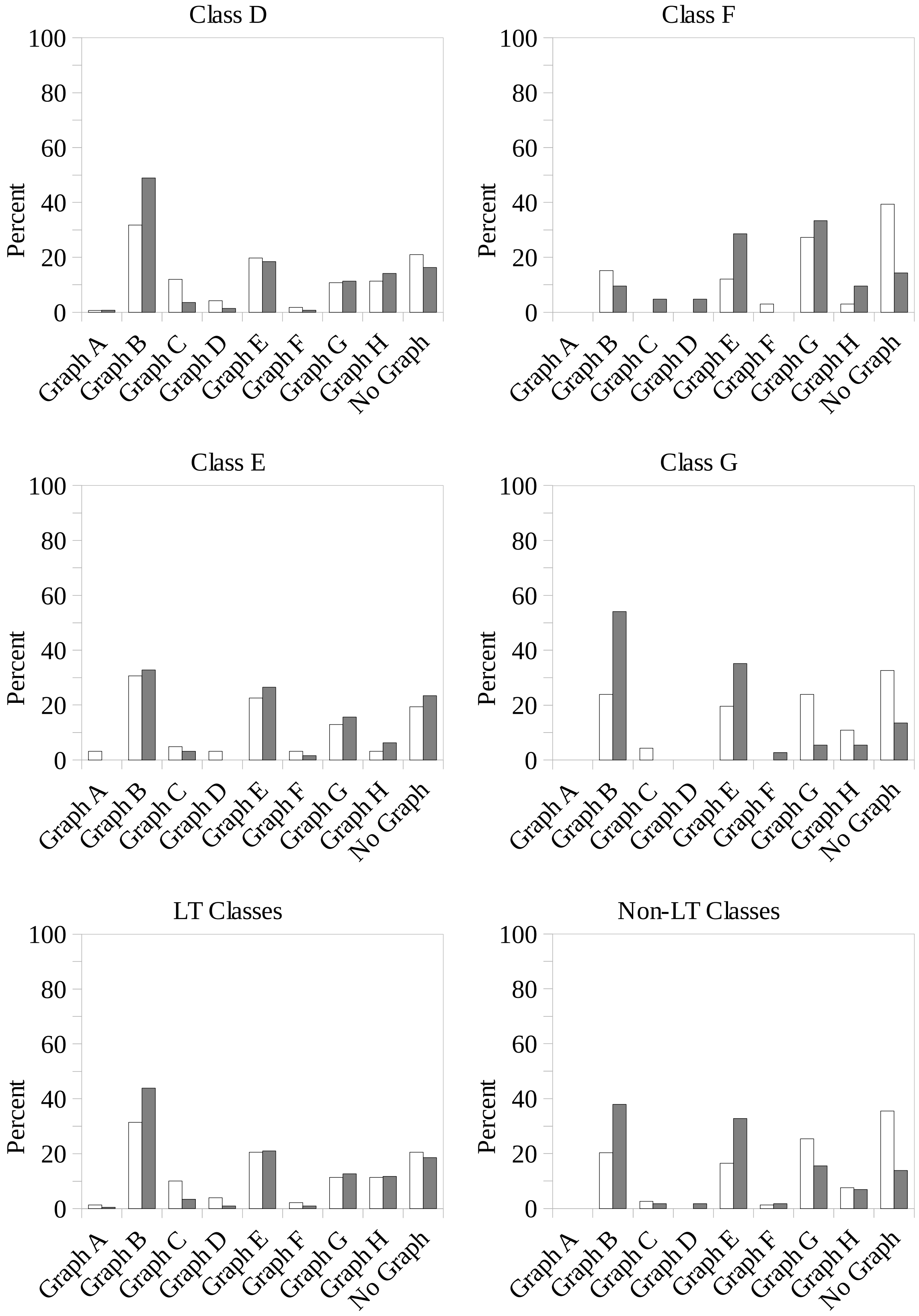


Figure 7.10: Students' graph choices for item 2 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

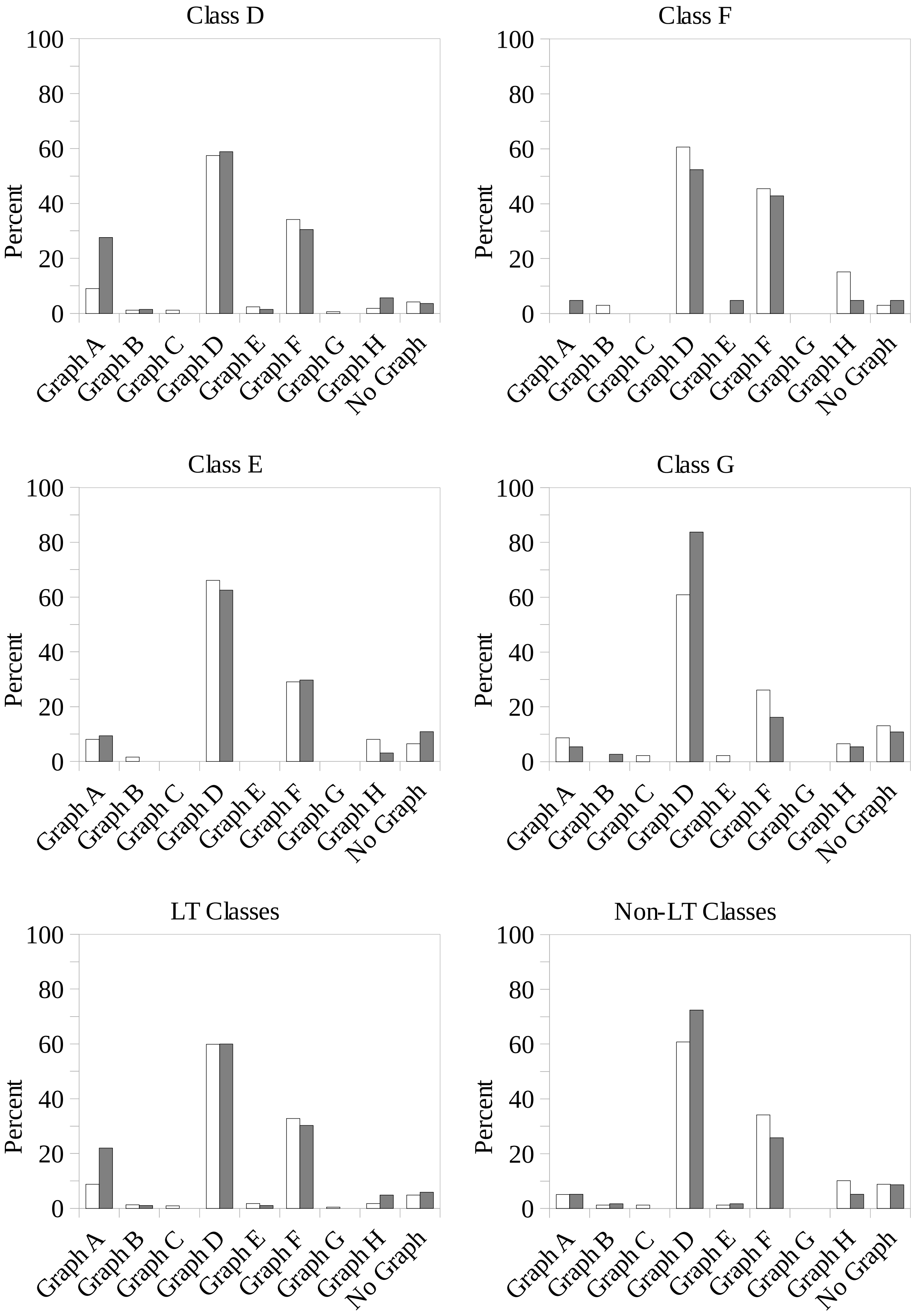


Figure 7.11: Students' graph choices for item 3 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

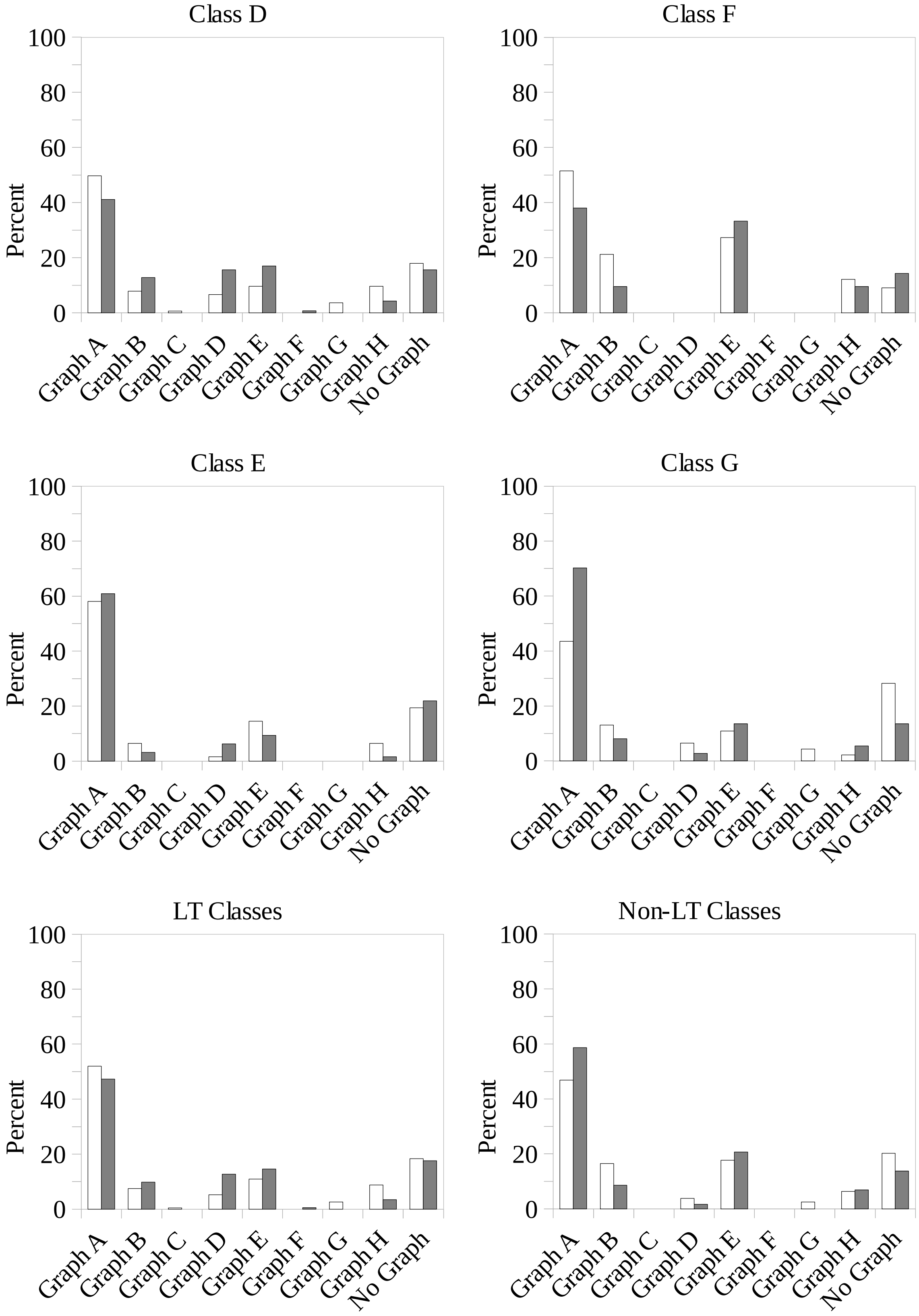


Figure 7.12: Students' graph choices for item 4 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Figures 7.9-7.12 show students' graph choices for items 1-4, respectively. These figures show that, for each item, students are gravitating toward the same graphs as they did in the fall 2009. An examination of the reasons students give for their graph selections reveals they are the same as in the fall 2009. In most cases, students either focus on the wrong feature of a graph or fail to construct complete and correct reasons for choosing a particular graph.

Figure 7.1 showed non-trivial normalized gains for both lecture-tutorial and non-lecture-tutorial students on item 5. This makes sense in light on Tables 7.12 and 7.13 which show the percent of students in score category 2 increasing pre- to post-instruction for both populations. For the lecture-tutorial classes, this percent increased by 23%, while for the non-lecture-tutorial classes it increased by 12%. What are these students actually saying in their answers?

The various reasoning elements used by students to answer this item are shown in Table 7.14 (pre-instruction) and Table 7.15 (post-instruction). The scoring rubric in Appendix F states that students simply need to draw or discuss that the slope of a Hubble plot should become steeper for a universe with a faster expansion rate. According to Table 7.14, a little over 60% of students in both populations give such a response pre-instruction. Post-instruction, this increases to 69% for the non-lecture-tutorial students and 82% for the lecture-tutorial students. About 40% of students in both populations, pre- and post-instruction, also mention an "increased velocity" in their responses. The response patterns in Tables 7.14 and 7.15 show that many students can give the correct answer to this item for the correct reason.

At the end of Chapter 6, I described how we added item 5 to Form A with the intent that it would be an easier item than items 1-4, and thus help use better estimate the abilities of low ability students. I also described how we made some changes to the "Hubble's Law" lecture-tutorial to further stress the relationship between the slope of a Hubble plot and the expansion rate of the universe. The response patterns for item 5 suggest that these revisions may have had some effect.

The distribution of scores in Tables 7.12 and 7.13 for item 6 also show some evidence of improvement. Pre-instruction, a high percent of students earned scores of 0: 11% for the lecture-tutorial students and 24% for the non-lecture-tutorial students. Post-instruction, these dropped

Table 7.14: Common reasoning elements used by students (pre-instruction) in their answers to Form A, Item 5: Use the blank graph provided below to draw what your think Figure 1 would look like if the universe had been expanding twice as fast. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student draws or discusses a steeper slope.	57%	79%	63%	76%	50%	61%
Student draws or discusses a flatter slope.	6%	5%	6%	9%	7%	8%
Student draws or discusses an unchanged slope.	15%	5%	12%	12%	15%	14%
Student draws a line with a non-constant slope.	8%	5%	7%	3%	4%	4%
Student draws a line whose slope is ≤ 0.	%	2%	0%	3%	0%	1%
Student draws a line whose y-intercept $\neq 0$.	%	3%	8%	0%	9%	5%
Student talks about increased velocity.	45%	37%	43%	52%	39%	44%
Student talks about decreased velocity.	0%	0%	0%	0%	0%	0%
Student talks about velocity staying the same.	1%	2%	1%	0%	2%	1%
Student talks about increased distance.	16%	23%	18%	15%	13%	14%
Student talks about decreased distance.	5%	0%	3%	0%	0%	0%
Student talks about distance staying the same.	2%	2%	2%	15%	2%	8%
Student compares variable x to variable y by saying x is more than y.	7%	18%	10%	9%	0%	4%
Student says distance and velocity change by the same amount.	3%	3%	3%	6%	2%	4%
Student says s/he can't answer the question without a time axis/variable.	2%	3%	2%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	15%	0%	11%	3%	4%	4%
Student has no idea.	4%	2%	3%	6%	4%	5%
Answer field is blank.	2%	0%	1%	0%	17%	10%

Table 7.15: Common reasoning elements used by students (post-instruction) in their answers to Form A, Item 5: Use the blank graph provided below to draw what your think Figure 1 would look like if the universe had been expanding twice as fast. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student draws or discusses a steeper slope.	82%	83%	82%	76%	65%	69%
Student draws or discusses a flatter slope.	6%	5%	5%	10%	3%	5%
Student draws or discusses an unchanged slope.	13%	5%	3%	14%	14%	14%
Student draws a line with a non-constant slope.	5%	3%	4%	0%	14%	9%
Student draws a line whose slope is ≤ 0.0	0%	0%	0%	0%	0%	0%
Student draws a line whose y-intercept $\neq 0.10$	3%	2%	2%	5%	5%	5%
Student talks about increased velocity.	37%	44%	39%	48%	38%	41%
Student talks about decreased velocity.	1%	0%	0%	0%	0%	0%
Student talks about velocity staying the same.	0%	0%	0%	5%	0%	2%
Student talks about increased distance.	9%	28%	15%	19%	19%	19%
Student talks about decreased distance.	1%	0%	1%	0%	5%	3%
Student talks about distance staying the same.	4%	0%	2%	0%	3%	2%
Student compares variable x to variable y by saying x is more than y.	2%	23%	9%	5%	5%	5%
Student says distance and velocity change by the same amount.	0%	3%	1%	0%	0%	0%
Student says s/he can't answer the question without a time axis/variable.	0%	2%	0%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	4%	5%	4%	10%	3%	5%
Student has no idea.	1%	0%	0%	0%	0%	0%
Answer field is blank.	1%	3%	1%	0%	3%	2%

Table 7.16: Common reasoning elements used by students (pre-instruction) in their answers to Form A, Item 6: Use the blank graph provided below to draw what your think Figure 1 would look like for a much older universe. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student draws or discusses a positive slope.	57%	50%	55%	64%	41%	51%
Student draws or discusses a negative slope.	7%	10%	7%	3%	7%	5%
Student draws or discusses a slope of zero.	4%	8%	5%	6%	4%	5%
Student draws or discusses a non-constant slope.	23%	21%	23%	21%	13%	16%
Student draws or discusses a steeper slope.	3%	0%	2%	0%	9%	5%
Student draws or discusses a flatter slope.	17%	21%	18%	27%	17%	22%
Student draws or discusses an unchanged slope.	17%	15%	17%	18%	13%	15%
Student draws a line whose y-intercept $\neq 0$.	14%	16%	14%	3%	9%	6%
Student talks about a faster expansion rate.	0%	0%	0%	0%	0%	0%
Student talks about a slower expansion rate.	8%	8%	8%	18%	0%	8%
Student talks about an unchanged expansion rate (i.e. expansion rate is the same as Figure 1).	2%	0%	2%	6%	0%	3%
Student talks about an expansion rate that speeds up.	1%	2%	1%	3%	0%	1%
Student talks about an expansion rate that slows down.	7%	10%	7%	0%	7%	4%
Student talks about a constant expansion rate (i.e. expansion rate doesn't change over time).	3%	8%	4%	0%	0%	0%
Student talks about an increased velocity/speed.	6%	0%	4%	0%	7%	4%
Student talks about a decreased velocity/speed.	18%	8%	15%	9%	9%	9%
Student talks about a constant velocity/speed.	3%	2%	3%	12%	2%	6%
Student talks about an increased distance.	5%	3%	5%	9%	13%	11%
Student talks about a decreased distance.	4%	0%	3%	3%	2%	3%
Student talks about a constant distance.	1%	0%	0%	0%	0%	0%
Student talks about what happened to the universe in the past.	1%	3%	2%	3%	0%	1%
Student talks about what will happen to the universe in the future.	5%	6%	6%	0%	2%	1%
Student says there's not enough information.	22%	24%	23%	27%	11%	18%
Students says the universe's age is irrelevant.	3%	6%	4%	6%	0%	3%
Student talks about the time the universe needs to reach its current size.	1%	0%	1%	0%	0%	0%
Student says s/he needs to know how the expansion rate changes.	11%	11%	11%	18%	0%	8%
Student says there's no time variable/axis.	2%	2%	2%	0%	0%	0%
Student says the significance of the universe's age is unclear.	7%	5%	7%	9%	4%	6%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	10%	18%	12%	3%	11%	6%
Student has no idea.	10%	5%	9%	3%	9%	20%
Answer field is blank.	6%	6%	6%	6%	30%	3%

Table 7.17: Common reasoning elements used by students (post-instruction) in their answers to Form A, Item 6: Use the blank graph provided below to draw what your think Figure 1 would look like for a much older universe. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student draws or discusses a positive slope.	88%	84%	87%	57%	78%	71%
Student draws or discusses a negative slope.	4%	0%	2%	5%	5%	5%
Student draws or discusses a slope of zero.	0%	0%	0%	5%	0%	2%
Student draws or discusses a non-constant slope.	21%	17%	20%	14%	16%	16%
Student draws or discusses a steeper slope.	6%	2%	4%	10%	8%	9%
Student draws or discusses a flatter slope.	52%	41%	49%	10%	32%	24%
Student draws or discusses an unchanged slope.	9%	20%	13%	24%	24%	24%
Student draws a line whose y-intercept $\neq 0$.	4%	5%	4%	5%	3%	3%
Student talks about a faster expansion rate.	1%	2%	1%	5%	0%	2%
Student talks about a slower expansion rate.	18%	28%	21%	0%	11%	7%
Student talks about an unchanged expansion rate (i.e. expansion rate is the same as Figure 1).	1%	0%	0%	5%	0%	2%
Student talks about an expansion rate that speeds up.	9%	6%	8%	0%	0%	0%
Student talks about an expansion rate that slows down.	9%	2%	6%	0%	5%	3%
Student talks about a constant expansion rate (i.e. expansion rate doesn't change over time).	1%	6%	2%	0%	5%	3%
Student talks about an increased velocity/speed.	2%	5%	3%	14%	5%	9%
Student talks about a decreased velocity/speed.	15%	13%	14%	5%	5%	5%
Student talks about a constant velocity/speed.	2%	5%	3%	10%	0%	3%
Student talks about an increased distance.	9%	8%	9%	19%	19%	19%
Student talks about a decreased distance.	1%	0%	0%	5%	0%	2%
Student talks about a constant distance.	0%	2%	0%	0%	0%	0%
Student talks about what happened to the universe in the past.	4%	3%	3%	5%	0%	2%
Student talks about what will happen to the universe in the future.	4%	0%	2%	0%	0%	0%
Student says there's not enough information.	7%	11%	8%	33%	8%	17%
Students says the universe's age is irrelevant.	1%	2%	1%	0%	0%	0%
Student talks about the time the universe needs to reach its current size.	9%	6%	8%	0%	5%	3%
Student says s/he needs to know how the expansion rate changes.	3%	3%	3%	14%	5%	9%
Student says there's no time variable/axis.	0%	5%	1%	10%	0%	3%
Student says the significance of the universe's age is unclear.	0%	0%	0%	5%	0%	2%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	9%	13%	10%	5%	5%	5%
Student has no idea.	6%	3%	5%	5%	5%	5%
Answer field is blank.	1%	3%	2%	0%	3%	2%

to 5% and 9%, respectively, while the percent of students with a score of 2 increased by 36% for the lecture-tutorial students and 8% for the non-lecture-tutorial students. However, few students earned the maximum score of 3. Pre-instruction, 0% of students received a 3 and post-instruction this only increased for the lecture-tutorial students and only to 3%.

In order to receive full credit on item 6, a student had to draw a line with a flatter slope and defend this drawing by saying that in order for the universe to be older it would need to take longer to reach its current size (or, more properly, a given region of the universe needs longer to reach its current size – but that is not a detail on which we dwell), which implies a slower expansion rate, which implies a flatter slope. Tables 7.16 and 7.17 show the reasoning elements used by students in response to this item pre- and post-instruction, respectively. According to Table 7.16, about 20% of both populations draw or discuss a flatter slope pre-instruction, while approximately another 20% draw or discuss a slope that is not constant. Post-instruction, the values continue to hover around 20% for the non-lecture-tutorial students, while 49% of lecture-tutorial students draw or discuss a flatter slope. Pre-instruction, 8% of both lecture-tutorial and non-lecture-tutorial students talk about a slower expansion rate. Post-instruction, these percents are 21% and 7%, respectively. How many students talk about the time the universe needs to reach its size? Only 1% (8%) of lecture-tutorial students pre-instruction (post-instruction) and 0% (3%) of non-lecture-tutorial students pre-instruction (post-instruction). Thus, while the lecture-tutorial students use more of the correct reasoning elements post-instruction than the non-lecture-tutorial students, neither group has a high percentage of students talking about the time the universe needs to reach a particular size.

7.4.2 Form B Responses

Tables 7.18 and 7.19 show the distribution of scores on Form B pre- and post-instruction, respectively. Both populations of students show shifts in the distribution of scores for most items pre- to post-instruction. An examination of Tables 7.18 and 7.19 reveals items for which the lecture-tutorial students appear to exhibit greater improvement than the non-lecture-tutorial students (e.g. item 2), items for which the non-lecture-tutorial students appear to exhibit greater improvement

Table 7.18: Pre-instruction distribution of scores on Form B for spring 2010.

	Class D					Class E					LT Classes				
	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Item 1	3%	57%	39%	1%	1%	12%	28%	51%	0%	9%	5%	50%	42%	0%	3%
Item 2	7%	33%	58%	2%	-	2%	23%	74%	2%	-	6%	31%	62%	2%	-
Item 3	3%	33%	64%	-	-	3%	34%	63%	-	-	3%	33%	64%	-	-
Item 4	2%	88%	10%	1%	-	0%	88%	12%	0%	-	1%	88%	11%	0%	-
Item 5	5%	76%	6%	12%	-	8%	58%	11%	23%	-	6%	72%	8%	15%	-
Item 6	14%	26%	18%	41%	-	12%	36%	18%	34%	-	14%	29%	18%	39%	-
Item 7	6%	64%	30%	-	-	3%	54%	43%	-	-	5%	62%	33%	-	-
	Class F					Class G					Non-LT Classes				
	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Item 1	0%	15%	73%	8%	4%	19%	6%	69%	0%	6%	10%	10%	71%	4%	5%
Item 2	0%	35%	58%	8%	-	9%	13%	75%	3%	-	5%	22%	67%	5%	-
Item 3	4%	23%	73%	-	-	9%	34%	56%	-	-	7%	29%	64%	-	-
Item 4	0%	81%	15%	4%	-	9%	88%	3%	0%	-	5%	84%	9%	2%	-
Item 5	0%	58%	27%	15%	-	6%	81%	0%	13%	-	3%	71%	12%	14%	-
Item 6	8%	65%	15%	12%	-	25%	38%	13%	25%	-	17%	50%	14%	19%	-
Item 7	4%	65%	31%	-	-	9%	63%	28%	-	-	7%	64%	29%	-	-

Table 7.19: Post-instruction distribution of scores on Form B for spring 2010.

	Class D					Class E					LT Classes				
	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Item 1	1%	1%	72%	12%	14%	2%	3%	48%	0%	48%	1%	2%	60%	9%	18%
Item 2	1%	1%	29%	69%	-	3%	5%	67%	25%	-	1%	1%	34%	54%	-
Item 3	6%	28%	67%	-	-	5%	16%	79%	-	-	5%	23%	61%	-	-
Item 4	2%	46%	46%	6%	-	5%	51%	31%	13%	-	2%	40%	41%	7%	-
Item 5	6%	40%	24%	29%	-	7%	49%	13%	31%	-	5%	36%	21%	28%	-
Item 6	14%	21%	10%	56%	-	20%	21%	7%	52%	-	14%	19%	8%	49%	-
Item 7	6%	42%	52%	-	-	11%	20%	69%	-	-	7%	34%	49%	-	-
	Class F					Class G					Non-LT Classes				
	0	1	2	3	4	0	1	2	3	4	0	1	2	3	4
Item 1	0%	19%	52%	5%	24%	5%	5%	32%	3%	55%	3%	10%	39%	3%	44%
Item 2	10%	10%	76%	5%	-	5%	8%	76%	11%	-	7%	8%	76%	8%	-
Item 3	5%	14%	81%	-	-	3%	39%	58%	-	-	3%	31%	66%	-	-
Item 4	10%	62%	14%	14%	-	5%	79%	16%	0%	-	7%	73%	15%	5%	-
Item 5	5%	62%	24%	10%	-	0%	68%	3%	29%	-	2%	66%	10%	22%	-
Item 6	0%	33%	29%	38%	-	16%	47%	8%	29%	-	10%	42%	15%	32%	-
Item 7	0%	38%	62%	-	-	0%	39%	61%	-	-	0%	39%	61%	-	-

(e.g. item 7), and even an item for which neither group demonstrates much improvement (item 3). What are students actually saying in their responses to these items?

Tables 7.20 and 7.21 show the percent of students who use each of the listed reasoning elements in their responses to item 1. The overall pattern of responses in Tables 7.20 and 7.21 is similar to those observed in the fall 2009 data for item 1, which makes sense given that we did not change the wording of this item. Many students, pre- and post-instruction, are describing the expansion of the universe in terms of the size of the universe increasing over time. However, a significant percent of students also talk about expansion as a metaphor for the increase in our knowledge over time and/or for the formation of new objects in the universe over time. Interesting, the former claim is more common in the pre-instruction lecture-tutorial students and the latter is more common among the post-instruction non-lecture-tutorial students. We have no hypothesis on why this should be so.

The responses to item 2 also tell a very similar story as our fall 2009 data. Tables 7.22 and 7.23 give the percent of reasoning elements used by both lecture-tutorial and non-lecture-tutorial students, pre- and post-instruction. Once again, only about 10% of students in both groups connected the Big Bang to the expansion of the universe in the pre-instruction responses. Post-instruction, this increases to 68% for the lecture-tutorial students and to just 15% for the non-lecture-tutorial students. 50% or more of all students pre-instruction say the Big Bang was an explosion. This drops to 12% for the lecture-tutorial students, but stays above 50% for the non-lecture-tutorial students. Finally, around 30% of all students pre-instruction claim that matter existed before the Big Bang. At the end of the semester only 6% of the lecture-tutorial students made such a claim, compared to 25% of the non-lecture-tutorial students. These responses are consistent with the idea that the lecture-tutorials help students develop more expert-like conceptions about the Big Bang.

We added item 3 to Form B for the first time for the spring 2010 semester. As the P-values, item step difficulties, and Thurstonian thresholds show, this is the easiest item for students on Form B. Tables 7.18 and 7.19 also show that the distribution of scores do not change much on this

Table 7.20: Common reasoning elements used by students (pre-instruction) in their answers to Form B, Item 1: Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding."

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says the size of the universe increases over time.	30%	31%	31%	35%	44%	40%
Student says the universe has a center.	6%	8%	7%	12%	3%	7%
Student says the universe has no center.	1%	0%	0%	4%	0%	2%
Student says the universe has an edge.	1%	0%	1%	15%	3%	9%
Student says the universe has no edge.	6%	0%	5%	19%	0%	9%
Student talks about redshifts/Doppler shifts.	1%	3%	1%	0%	0%	0%
Student says space(time) is growing/stretching.	1%	5%	2%	12%	0%	5%
Student talks about the movement of galaxies and/or their increasing distances.	3%	18%	7%	8%	16%	12%
Student talks about the movement of stars and/or their increasing distances.	5%	11%	6%	12%	9%	10%
Student talks about the movement of planets and/or their increasing distances.	4%	3%	4%	8%	3%	5%
Student talks about the movement of "objects" (something unspecified or not a star, planet, or galaxy) and/or their increasing distances.	10%	5%	8%	12%	19%	16%
Student says the distances between everything increase.	5%	8%	6%	35%	9%	21%
Student says farther objects move away faster.	0%	5%	1%	0%	0%	0%
Student talks about the Big Bang.	13%	15%	13%	27%	9%	17%
Student talks about an explosion.	4%	3%	4%	4%	6%	5%
Student says the early universe was once hot, small, and/or dense.	0%	8%	2%	0%	0%	0%
Student says we learn more about the universe over time.	40%	22%	35%	8%	6%	7%
Student talks about how we are looking further back in time as we look farther in space.	0%	0%	0%	8%	0%	3%
Student says new things are created in the universe over time.	19%	11%	17%	4%	0%	2%
Student gives irrelevant information.	0%	3%	1%	0%	0%	0%
Student gives some other reason not specified above.	12%	6%	10%	0%	6%	3%
Student says s/he has no idea.	1%	2%	1%	0%	6%	3%
Answer field is blank or the student provided no reason.	1%	6%	2%	0%	13%	7%

Table 7.21: Common reasoning elements used by students (post-instruction) in their answers to Form B, Item 1: Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding."

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says the size of the universe increases over time.	80%	20%	63%	24%	13%	17%
Student says the universe has a center.	1%	0%	0%	5%	5%	5%
Student says the universe has no center.	2%	2%	2%	0%	0%	0%
Student says the universe has an edge.	0%	0%	0%	5%	5%	5%
Student says the universe has no edge.	1%	0%	1%	5%	0%	2%
Student talks about redshifts/Doppler shifts.	5%	5%	5%	14%	11%	12%
Student says space(time) is growing/stretching.	21%	38%	26%	0%	13%	8%
Student talks about the movement of galaxies and/or their increasing distances.	18%	59%	30%	38%	63%	54%
Student talks about the movement of stars and/or their increasing distances.	3%	7%	4%	5%	8%	7%
Student talks about the movement of planets and/or their increasing distances.	2%	3%	2%	0%	3%	2%
Student talks about the movement of "objects" (something unspecified or not a star, planet, or galaxy) and/or their increasing distances.	12%	13%	12%	10%	3%	5%
Student says the distances between everything increase.	20%	10%	17%	19%	11%	14%
Student says farther objects move away faster.	2%	13%	5%	5%	3%	3%
Student talks about the Big Bang.	0%	5%	1%	19%	11%	14%
Student talks about an explosion.	12%	0%	9%	0%	3%	2%
Student says the early universe was once hot, small, and/or dense.	12%	3%	10%	0%	0%	0%
Student says we learn more about the universe over time.	1%	2%	1%	0%	0%	0%
Student talks about how we are looking further back in time as we look farther in space.	0%	2%	0%	0%	3%	2%
Student says new things are created in the universe over time.	2%	0%	1%	10%	13%	12%
Student gives irrelevant information.	0%	2%	0%	0%	0%	0%
Student gives some other reason not specified above.	1%	5%	2%	14%	8%	10%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%
Answer field is blank or the student provided no reason.	0%	0%	0%	0%	0%	0%

Table 7.22: Common reasoning elements used by students (pre-instruction) in their answers to Form B, Item 2: Explain, in as much detail as possible, what astronomers mean by the “Big Bang Theory.”

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says the Big Bang is the beginning of the universe.	43%	49%	44%	35%	66%	52%
Student says the Big Bang is the beginning of expansion.	11%	8%	10%	8%	9%	9%
Student says the Big Bang was the beginning of something smaller than the universe.	23%	23%	23%	31%	9%	19%
Student says the Big Bang is an event that happened to something smaller than the universe.	16%	0%	12%	0%	3%	2%
Student says the Big Bang is the beginning of space.	1%	0%	0%	0%	0%	0%
Student says the Big Bang is the beginning of time.	1%	2%	1%	0%	3%	2%
Student talks about the creation/production of elements.	6%	9%	7%	15%	0%	7%
Student says the Big Bang was an explosion.	51%	48%	50%	62%	56%	59%
Student says the Big Bang was not an explosion.	0%	2%	0%	0%	0%	0%
Student says matter existed before the Big Bang.	26%	32%	27%	46%	22%	33%
Student says there was a dense piece of matter before the Big Bang.	7%	11%	8%	12%	3%	7%
Student talks about matter coming together before the Big Bang.	5%	6%	6%	15%	0%	7%
Student says matter formed from energy.	2%	0%	2%	0%	0%	0%
Student says the early universe was hot, dense, and/or small.	0%	3%	1%	4%	0%	2%
Student says the Big Bang was an event that happened in empty space.	10%	2%	8%	8%	9%	9%
Student gives irrelevant information.	0%	2%	0%	0%	0%	0%
Student gives some other reason not specified above.	9%	8%	9%	0%	0%	0%
Student says s/he has no idea.	5%	2%	4%	0%	0%	0%
Answer field is blank or the student provided no reason.	2%	0%	1%	0%	9%	5%

Table 7.23: Common reasoning elements used by students (post-instruction) in their answers to Form B, Item 2: Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory."

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says the Big Bang is the beginning of the universe.	46%	49%	47%	29%	71%	56%
Student says the Big Bang is the beginning of expansion.	80%	38%	68%	14%	16%	15%
Student says the Big Bang was the beginning of something smaller than the universe.	2%	5%	3%	10%	8%	8%
Student says the Big Bang is an event that happened to something smaller than the universe.	0%	0%	0%	0%	0%	0%
Student says the Big Bang is the beginning of space.	5%	0%	4%	5%	0%	2%
Student says the Big Bang is the beginning of time.	7%	0%	5%	5%	0%	2%
Student talks about the creation/production of elements.	1%	11%	4%	14%	8%	10%
Student says the Big Bang was an explosion.	6%	26%	12%	62%	55%	58%
Student says the Big Bang was not an explosion.	4%	0%	3%	0%	0%	0%
Student says matter existed before the Big Bang.	3%	15%	6%	29%	24%	25%
Student says there was a dense piece of matter before the Big Bang.	1%	7%	3%	10%	5%	7%
Student talks about matter coming together before the Big Bang.	0%	2%	0%	5%	5%	5%
Student says matter formed from energy.	32%	18%	28%	0%	0%	0%
Student says the early universe was hot, dense, and/or small.	0%	31%	9%	0%	16%	10%
Student says the Big Bang was an event that happened in empty space.	0%	0%	0%	5%	5%	5%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	4%	7%	5%	10%	3%	5%
Student says s/he has no idea.	0%	0%	0%	0%	3%	2%
Answer field is blank or the student provided no reason.	1%	2%	1%	10%	3%	5%

Table 7.24: Common reasoning elements used by students (pre-instruction) in their answers to Form B, Item 3: Each dot in the picture on the left is a galaxy. The Milky Way Galaxy (the one we live in) is at the center of the picture. All of the galaxies inside the circle can be seen from Earth. Any galaxies that exist outside the circle are so far away that their light has not had time to reach Earth. Describe what inhabitants of Galaxy X probably see when they look in the direction of the arrow.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says they'd see more galaxies.	51%	62%	54%	65%	53%	59%
Student says they'd see more stars.	9%	11%	9%	8%	6%	7%
Student says they'd see more planets.	4%	3%	4%	0%	0%	0%
Student says they'd see more "objects" (otherwise unspecified).	4%	3%	4%	0%	0%	0%
Student says they'd see something similar to what we see.	31%	23%	29%	19%	25%	22%
Student says they'd see things we cannot see.	25%	25%	25%	38%	16%	26%
Student says they'd see nothing/blackness/empty space.	22%	25%	23%	15%	25%	21%
Student says they'd see objects separated by greater distances than what we see.	0%	3%	1%	4%	0%	2%
Student says Galaxy X's observable universe extends into regions outside of our observable universe.	34%	14%	29%	42%	19%	29%
Student gives irrelevant information.	1%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	7%	3%	6%	4%	13%	9%
Student says s/he has no idea.	1%	0%	1%	0%	0%	0%
Response field is blank.	2%	3%	2%	4%	9%	7%

Table 7.25: Common reasoning elements used by students (post-instruction) in their answers to Form B, Item 3: Each dot in the picture on the left is a galaxy. The Milky Way Galaxy (the one we live in) is at the center of the picture. All of the galaxies inside the circle can be seen from Earth. Any galaxies that exist outside the circle are so far away that their light has not had time to reach Earth. Describe what inhabitants of Galaxy X probably see when they look in the direction of the arrow.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says they'd see more galaxies.	49%	48%	48%	52%	47%	49%
Student says they'd see more stars.	7%	5%	6%	19%	3%	8%
Student says they'd see more planets.	1%	0%	1%	5%	0%	2%
Student says they'd see more "objects" (otherwise unspecified).	3%	0%	2%	5%	5%	5%
Student says they'd see something similar to what we see.	30%	48%	35%	43%	29%	34%
Student says they'd see things we cannot see.	32%	18%	28%	29%	18%	22%
Student says they'd see nothing/blackness/empty space.	17%	5%	13%	14%	21%	19%
Student says they'd see objects separated by greater distances than what we see.	1%	0%	0%	0%	0%	0%
Student says Galaxy X's observable universe extends into regions outside of our observable universe.	44%	25%	39%	24%	21%	22%
Student gives irrelevant information.	3%	2%	2%	0%	0%	0%
Student gives some other reason not specified above.	5%	3%	5%	0%	5%	3%
Student says s/he has no idea.	2%	0%	1%	0%	3%	2%
Response field is blank.	1%	3%	1%	5%	0%	2%

item. The most common reasoning elements invoked by students responding to this item along with the percent of students using each reasoning element are given in Tables 7.24 and 7.19. The most popular reasoning elements pre- and post-instruction are correct or part of correct answers. However, a significant minority of students pre- and post-instruction maintain that a galaxy at the edge of our observable universe will see nothing but blackness if they look 180^o away from the Milky Way. Here are three typical responses that falls into this category:

> "Galaxy X inhabitants probably can't see anything because their galaxy has not yet received light."

> "They probably see nothing since there is no light able to reach there."

> "They see other galaxies, and at the edge they see dark matter, or areas without any light."

These are the kinds of responses one might expect from students who have not accepted the cosmological principle.

Figure 7.13 shows the percent of students who select each of the answer choices to item 4. There is a definite shift of lecture-tutorial students away from one of the incorrect answers (B) toward the correct answer (C) pre- to post-instruction. This shift is smaller for the non-lecture-tutorial students. Tables 7.26-7.29 give the reasoning elements used by students in their responses pre- and post-instruction. These tables show that many students who select choices A or B frequently talk about matter and/or space existing before the Big Bang. This is consistent with the prior findings of Prather, Slater, and Offerdahl (2003) as well as students' responses to other items in this study (especially item 2 on this survey). Students who select choice C frequently state that there was no matter and/or space before the Big Bang. The fact that a greater percent of lecture-tutorial students select option C suggests that the lecture-tutorials are helping students develop a more accurate understanding of the nature of the Big Bang.

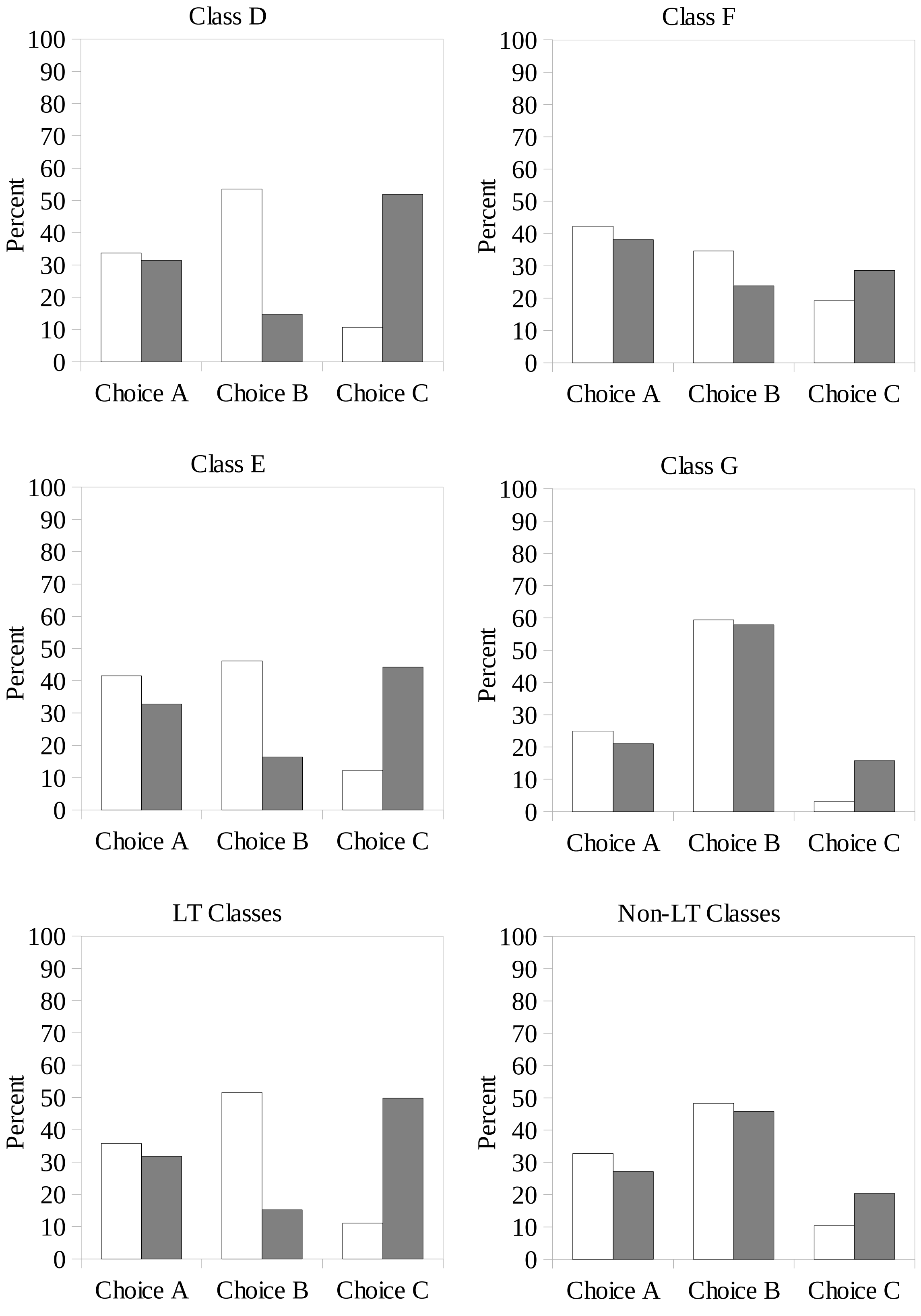


Figure 7.13: Percent of students who selected choice A, B, or C in response to item 4 on Form B. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Tables 7.30 and 7.31 give the percent of students who used the various listed reasoning elements in their responses to item 5. Either or both of the first two listed reasoning elements in these tables are considered correct. The first of these ("Student says that the universe is infinite/has no edges") was invoked by approximately 15% of respondents from both populations, pre- and post-instruction. The second ("Student says the universe is the same everywhere") was used by 0% of the lecture-tutorial students and 2% of the non-lecture-tutorial students pre-instruction. Post-instruction, these changed to 18% and 8%, respectively. The most common incorrect answer reflects the idea that the universe has a center and the center is where the Big Bang happened and/or where everything is expanding away from. This was stated by 21% of the lecture-tutorial students and 29% of the non-lecture-tutorial students pre-instruction. Post-instruction, these percents dropped to 12% and 15%, respectively. Another common incorrect answer states that the universe does not have a center or that the center of the universe changes because of expansion. Here is a typical student response that falls in this category:

> "No because the universe is constantly expanding in different directions so the 'center' changes."

The percent of students making such a claim actually increased pre- to post-instruction for both groups. This underscores the challenge of fostering expert-like understandings of the Big Bang and expansion of the universe in Astro 101 students.

Like items 5 and 6 on the fall 2009 version of Form B, items 6 and 7 on the spring 2010 version did not produce a sufficient diversity of responses to warrant creating detailed tables of students reasoning elements. According to the scoring rubric in Appendix F, students receive a 1 on item 6 if their answer implies that the distribution of galaxies peters out after traveling a certain distance. The distribution of scores in Tables 7.18 and 7.19 show that the percent of students earning a score of 1 decreased more for the lecture-tutorial students than for the non-lecture-tutorial students. The percent of students earning a score of 3 increased by a greater percent for the lecture-tutorial students than for the non-lecture-tutorial students. This is consistent with the hypothesis that the lecture-tutorials help students learn cosmology.

Table 7.26: Common reasoning elements used by lecture-tutorial students (pre-instruction) in their answers to Form B, Item 4: Circle the phrase that best completes the sentence. Surrounding the event called the BB was: a) a region of space that includes nothing (empty space); b) a region of space that includes particles and matter; c) neither a nor b. Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D			Class E			LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student says there was “nothing” (unspecified) before Big Bang.	21%	1%	30%	19%	0%	0%	20%	1%	21%
Student says there was “something” (unspecified) before the Big Bang.	2%	2%	0%	4%	3%	0%	2%	2%	0%
Student says there was no time before the Big Bang.	2%	0%	0%	0%	0%	13%	1%	0%	4%
Student says there was time before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was no space before the Big Bang.	6%	0%	10%	0%	0%	13%	4%	0%	11%
Student says there was space before the Big Bang.	22%	5%	15%	33%	3%	13%	26%	5%	14%
Student says there was no matter before the Big Bang.	13%	8%	10%	4%	0%	0%	10%	6%	7%
Student says there was matter before the Big Bang.	14%	66%	20%	15%	53%	38%	14%	63%	25%
Student says matter necessarily existed before the Big Bang.	0%	22%	0%	0%	30%	0%	0%	24%	0%
Students says matter before the Big Bang was all in one place/part of the Big Bang.	13%	2%	0%	11%	0%	25%	12%	2%	7%
Student says space is never completely empty.	2%	4%	0%	0%	10%	0%	1%	5%	0%
Student says there was only energy (no matter) at or before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0 %
Student says the Big Bang was the beginning of the universe/everything.	5%	4%	10%	19%	7%	0%	9%	5%	7 %
Student says the Big Bang is an event that happened to something smaller than the universe.	6%	15%	15%	0%	3%	13%	4%	12%	14%
Student says the Big Bang was an explosion.	10%	10%	10%	4%	13%	0%	8%	11%	7%
Student says we don’t/can’t know or describe what the universe was like this early.	2%	0%	5%	4%	0%	0%	2%	0%	4%
Student talks about what the universe was like after the Big Bang.	0%	7%	5%	0%	7%	0%	0%	7%	4%
Student says this answer makes sense/the alternatives don’t make sense.	0%	0%	0%	7%	13%	0%	2%	3%	0%
Student says this answer is better than the alternatives.	0%	0%	0%	0%	0%	13%	0%	0%	4%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	8%	1%	20%	7%	0%	13%	8%	1%	18%
Student has no idea.	10%	9%	10%	4%	7%	13%	8%	8%	11%
Answer field is blank.	21%	16%	5%	19%	23%	13%	20%	18%	7%

Table 7.27: Common reasoning elements used by non-lecture-tutorial students (pre-instruction) in their answers to Form B, Item 4: Circle the phrase that best completes the sentence. Surrounding the event called the BB was: a) a region of space that includes nothing (empty space); b) a region of space that includes particles and matter; c) neither a nor b. Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F			Class G			Non-LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student says there was “nothing” (unspecified) before Big Bang.	0%	0%	0%	50%	5%	0%	21%	4%	0%
Student says there was “something” (unspecified) before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was no time before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was time before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was no space before the Big Bang.	9%	0%	40%	0%	0%	0%	5%	0%	33%
Student says there was space before the Big Bang.	36%	0%	20%	0%	5%	0%	21%	4%	17%
Student says there was no matter before the Big Bang.	18%	0%	40%	38%	5%	0%	26%	4%	33%
Student says there was matter before the Big Bang.	45%	78%	20%	0%	47%	0%	26%	57%	17%
Student says matter necessarily existed before the Big Bang.	0%	44%	0%	0%	26%	0%	0%	32%	0%
Students says matter before the Big Bang was all in one place/part of the Big Bang.	36%	0%	0%	0%	5%	0%	21 %	4%	0%
Student says space is never completely empty.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was only energy (no matter) at or before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0 %
Student says the Big Bang was the beginning of the universe/everything.	9%	0%	20%	50%	0%	0%	26%	0%	17%
Student says the Big Bang is an event that happened to something smaller than the universe.	0%	0%	0%	0%	0%	0 %	0%	0%	0%
Student says the Big Bang was an explosion.	27%	11%	0%	0%	0%	0%	16%	4%	0%
Student says we don’t/can’t know or describe what the universe was like this early.	9%	0%	40%	0%	0%	0%	5%	0%	33%
Student talks about what the universe was like after the Big Bang.	9%	0%	0%	0%	0%	0%	5%	0%	0%
Student says this answer makes sense/the alternatives don’t make sense.	9%	0%	0%	0%	0%	0%	5%	0%	0%
Student says this answer is better than the alternatives.	0%	0%	0%	0%	5%	0%	0%	4%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	9%	0%	20%	13%	5%	0%	11%	4%	17%
Student has no idea.	0%	0%	0%	0%	11%	100%	0%	7%	17%
Answer field is blank.	9%	22%	0%	13%	26%	0%	11%	25%	0%

Table 7.28: Common reasoning elements used by lecture-tutorial students (post-instruction) in their answers to Form B, Item 4: Circle the phrase that best completes the sentence. Surrounding the event called the BB was: a) a region of space that includes nothing (empty space); b) a region of space that includes particles and matter; c) neither a nor b. Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D			Class E			LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student says there was "nothing" (unspecified) before Big Bang.	55%	4%	40%	15%	0%	4%	43%	3%	31%
Student says there was "something" (unspecified) before the Big Bang.	2%	0%	2%	0%	0%	7%	1%	0%	4%
Student says there was no time before the Big Bang.	0%	0%	6%	0%	0%	0%	0%	0%	5%
Student says there was time before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was no space before the Big Bang.	0%	0%	43%	0%	0%	52%	0%	0%	45%
Student says there was space before the Big Bang.	10%	0%	1%	10%	0%	0%	10%	0%	1%
Student says there was no matter before the Big Bang.	31%	9%	41%	35%	0%	26%	32%	6%	37%
Student says there was matter before the Big Bang.	0%	52%	2%	0%	0%	4%	0%	36%	3%
Student says matter necessarily existed before the Big Bang.	0%	17%	0%	0%	0%	0%	0%	12%	0%
Students says matter before the Big Bang was all in one place/part of the Big Bang.	0%	4%	0%	0%	0%	0%	0%	3%	0%
Student says space is never completely empty.	0%	4%	0%	0%	10%	0%	0%	6%	0%
Student says there was only energy (no matter) at or before the Big Bang.	0%	0%	0%	10%	10%	33%	3%	3%	8%
Student says the Big Bang was the beginning of the universe/everything.	0%	0%	0%	5%	0%	0%	1%	0%	0%
Student says the Big Bang is an event that happened to something smaller than the universe.	0%	4%	0%	0%	0%	0 %	0%	3%	0%
Student says the Big Bang was an explosion.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we don't/can't know or describe what the universe was like this early.	0%	0%	1%	0%	0%	15%	0%	0%	5%
Student talks about what the universe was like after the Big Bang.	2%	9%	0%	5%	40%	4%	3%	18%	1%
Student says this answer makes sense/the alternatives don't make sense.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says this answer is better than the alternatives.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	4%	9%	2%	5%	0%	7%	4%	6%	4%
Student has no idea.	2%	4%	0%	5%	0%	4%	3%	3%	1%
Answer field is blank.	12%	35%	2%	20%	60%	0%	14%	42%	2%

Table 7.29: Common reasoning elements used by non-lecture-tutorial students (post-instruction) in their answers to Form B, Item 4: Circle the phrase that best completes the sentence. Surrounding the event called the BB was: a) a region of space that includes nothing (empty space); b) a region of space that includes particles and matter; c) neither a nor b. Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F			Class G			Non-LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student says there was “nothing” (unspecified) before Big Bang.	13%	0%	0%	50%	0%	0%	31%	0%	0%
Student says there was “something” (unspecified) before the Big Bang.	0%	0%	17%	13%	0%	0%	6%	0%	8%
Student says there was no time before the Big Bang.	0%	0%	17%	0%	0%	0%	0%	0%	8%
Student says there was time before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there was no space before the Big Bang.	0%	0%	50%	0%	0%	17%	0%	0%	33%
Student says there was space before the Big Bang.	25%	0%	17%	13%	0%	0%	19%	0%	8%
Student says there was no matter before the Big Bang.	0%	0%	83%	13%	9%	17%	6%	7%	50%
Student says there was matter before the Big Bang.	25%	60%	0%	0%	45%	17%	13%	48%	8%
Student says matter necessarily existed before the Big Bang.	0%	40%	0%	0%	23%	0%	0%	26%	0%
Students says matter before the Big Bang was all in one place/part of the Big Bang.	25%	0%	0%	0%	0%	0%	13 %	0%	0%
Student says space is never completely empty.	0%	20%	0%	0%	0%	0%	0%	4%	0%
Student says there was only energy (no matter) at or before the Big Bang.	0%	0%	0%	0%	0%	0%	0%	0%	0 %
Student says the Big Bang was the beginning of the universe/everything.	25%	0%	0%	25%	9%	0%	25%	7%	0 %
Student says the Big Bang is an event that happened to something smaller than the universe.	0%	0%	0%	0%	0%	0 %	0%	0%	0%
Student says the Big Bang was an explosion.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we don’t/can’t know or describe what the universe was like this early.	13%	0%	17%	0%	5%	33%	6 %	4%	25%
Student talks about what the universe was like after the Big Bang.	0%	0%	0%	0%	9%	0%	0%	7%	0%
Student says this answer makes sense/the alternatives don’t make sense.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says this answer is better than the alternatives.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student has no idea.	0%	20%	0%	0%	0%	17%	0%	4%	8%
Answer field is blank.	25%	20%	0%	13%	41%	17%	19%	37%	8%

Table 7.30: Common reasoning elements used by students (pre-instruction) in their answers to Form B, Item 5: Independent of whether we know its true location, is there a center to the universe?

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says the universe is infinite/has no edges.	11%	23%	14%	15%	13%	14%
Student says the universe is the same everywhere.	1%	0%	0%	4%	0%	2%
Student says the Sun is at the center.	2%	2%	2%	0%	0%	0%
Student says everything has a center/must have a center.	15%	9%	13%	4%	16%	10%
Student says the center is where the Big Bang happened/where the universe began/where the universe is expanding from.	20%	23%	21%	38%	22%	29%
Student says galaxies orbit the center of the universe.	3%	5%	3%	8%	3%	5%
Student says our ignorance prevents us from seeing the center.	18%	6%	15%	8%	16%	12%
Student says technological limitations prevent us from seeing the center.	0%	0%	0%	0%	0%	0%
Student says there is no center or the center changes because of expansion/objects are in motion.	12%	5%	10%	23%	13%	17%
Student says there is no center since the universe has no shape/irregular shape.	1%	0%	1%	4%	0%	2%
Student gives irrelevant information.	1%	0%	1%	0%	0%	0%
Student gives some other reason not specified above.	14%	14%	14%	15%	19%	17%
Student says s/he has no idea.	4%	8%	5%	0%	0%	0%
Response field is blank.	5%	9%	6%	4%	6%	5%

Table 7.31: Common reasoning elements used by students (post-instruction) in their answers to Form B, Item 5: Independent of whether we know its true location, is there a center to the universe?

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says the universe is infinite/has no edges.	15%	18%	16%	5%	21%	15%
Student says the universe is the same everywhere.	17%	18%	18%	5%	11%	8%
Student says the Sun is at the center.	1%	0%	0%	0%	0%	0%
Student says everything has a center/must have a center.	2%	2%	2%	5%	3%	3%
Student says the center is where the Big Bang happened/where the universe began/where the universe is expanding from.	11%	16%	12%	24%	11%	15%
Student says galaxies orbit the center of the universe.	0%	2%	0%	0%	3%	2%
Student says our ignorance prevents us from seeing the center.	8%	10%	8%	24%	18%	20%
Student says technological limitations prevent us from seeing the center.	0%	2%	0%	10%	0%	3%
Student says there is no center or the center changes because of expansion/objects are in motion.	19%	26%	21%	19%	26%	24%
Student says there is no center since the universe has no shape/irregular shape.	4%	2%	4%	5%	3%	3%
Student gives irrelevant information.	1%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	14%	13%	14%	14%	18%	17%
Student says s/he has no idea.	1%	0%	1%	0%	0%	0%
Response field is blank.	0%	10%	3%	10%	3%	5%

What about item 7? Tables 7.18 and 7.19 reveal that while both populations show improvements pre- to post-instruction, these improvements are greater for the non-lecture-tutorial students. As noted in Chapter 6, this may be explained by the fact that the subject of item 7 is not emphasized by our cosmology lecture-tutorials.

7.4.3 Form C Responses

Tables 7.32 and 7.33 give the distribution of scores on Form C's items for the spring 2010. Like our data from the fall 2009, the data in Tables 7.32 and 7.33 show that the lecture-tutorial students generally make larger improvements than their non-lecture-tutorial peers. For most of Form C's items, the percent of students in lower scoring categories decreases by a greater percent, pre- to post-instruction, for the lecture-tutorial students than for the non-lecture-tutorial students. The percent of students in higher scoring categories increases by a greater amount, pre- to post-instruction, for the lecture-tutorial students than for the non-lecture-tutorial students. The exception to this trend is item 6, where both populations of students appear to migrate toward lower scores pre- to post-instruction.

Table 7.32: Pre-instruction distribution of scores on Form C for spring 2010.

	Class D				Class E				LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	1%	83%	10%	7%	7%	54%	20%	18%	3%	74%	13%	10%
Item 2	4%	57%	38%	1%	7%	70%	22%	1%	5%	61%	32%	1%
Item 3	7%	62%	20%	11%	8%	48%	6%	37%	8%	58%	15%	20%
Item 4	1%	75%	11%	13%	0%	66%	11%	23%	0%	72%	11%	16%
Item 5	2%	65%	19%	14%	2%	66%	11%	20%	2%	66%	16%	16%
Item 6	8%	47%	10%	35%	2%	40%	18%	40%	6%	44%	13%	36%
	Class F				Class G				Non-LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	8%	56%	13%	23%	2%	77%	7%	14%	5%	67%	10%	18%
Item 2	3%	74%	23%	0%	7%	58%	35%	0%	5%	66%	29%	0%
Item 3	8%	56%	10%	26%	21%	49%	9%	21%	15%	52%	10%	23%
Item 4	0%	85%	5%	10%	7%	72%	5%	16%	4%	78%	5%	13%
Item 5	0%	72%	8%	21%	5%	63%	19%	14%	2%	67%	13%	17%
Item 6	13%	26%	18%	44%	9%	42%	21%	28%	11%	34%	20%	35%

Table 7.33: Post-instruction distribution of scores on Form C for spring 2010.

	Class D				Class E				LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	0%	10%	11%	78%	0%	7%	15%	78%	0%	9%	13%	78%
Item 2	2%	20%	21%	57%	1%	47%	19%	32%	2%	29%	20%	49%
Item 3	1%	23%	9%	67%	1%	22%	4%	72%	1%	23%	7%	69%
Item 4	0%	50%	8%	42%	0%	43%	6%	51%	0%	48%	7%	45%
Item 5	1%	61%	11%	27%	6%	46%	6%	43%	2%	57%	9%	32%
Item 6	6%	51%	18%	25%	8%	40%	13%	39%	7%	47%	17%	30%
	Class F				Class G				Non-LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	0%	64%	18%	18%	3%	5%	30%	62%	2%	27%	25%	46%
Item 2	5%	82%	14%	0%	3%	32%	59%	5%	3%	51%	42%	3%
Item 3	0%	41%	23%	36%	3%	49%	5%	43%	2%	46%	12%	41%
Item 4	0%	68%	0%	32%	8%	62%	8%	22%	5%	64%	5%	25%
Item 5	0%	59%	23%	18%	5%	57%	19%	19%	3%	58%	20%	19%
Item 6	0%	41%	32%	27%	8%	30%	32%	30%	5%	34%	32%	29%

Overall, our spring 2010 data for Form C support our findings from our fall 2009 data. For example, Figure 7.14 shows the percent of students who say the temperature of the universe has increased, decreased, stayed the same, and changed (as before, I = the temperature increased, D = the temperature decreased, S = the temperature stayed the same, and C = the temperature changed). The results for the lecture-tutorial and non-lecture-tutorial classes are basically the same as they were in the fall 2009. Furthermore, the reasons students give for their answers also fall into the same categories delineated in Chapter 6, as Tables 7.34-7.37 show. Most of the students who claim the temperature has cooled support this claim by referring to the expansion of the universe. Many students who make incorrect statements do so for the same reasons previous discussed (e.g. the temperature has increased due to global warming). The similarities between the fall 2009 data and the spring 2010 data do not end with this item. The responses to item 2 (see Tables 7.38 and 7.39) and item 3 (see Figure 7.15 and Tables 7.40-7.43) likewise show similar response patterns and errors to their fall 2009 versions.

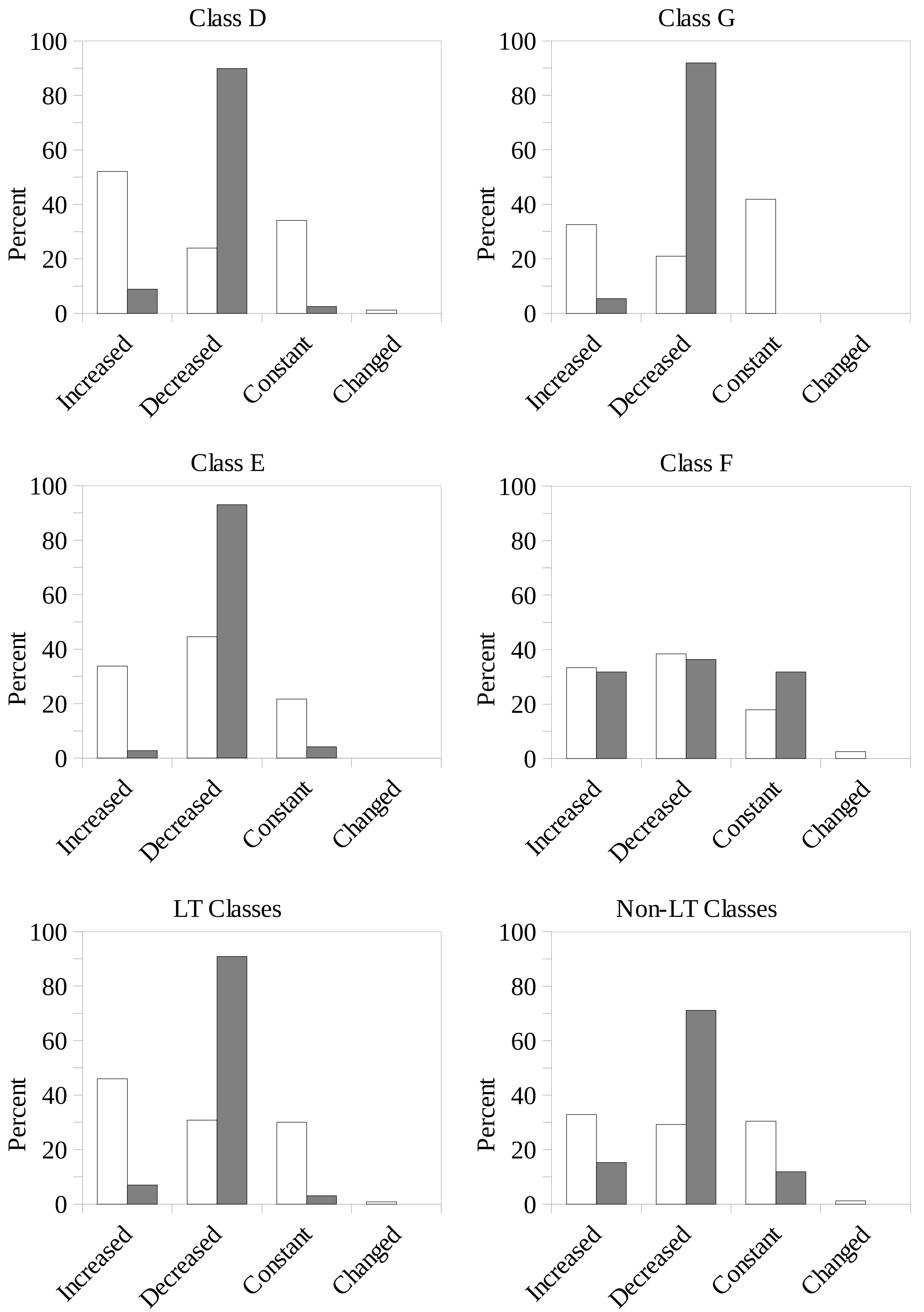


Figure 7.14: Percent of students who think the temperature of the universe has increased, decreased, stayed constant, or changed (without specifying if it went up or down). White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Table 7.34: Common reasoning elements used by lecture-tutorial students (pre-instruction) in their answers to Form C, Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	6%	25%	0%	0%	25%	57%	17%	0%	10%	40%	4%	0%
Student talks about the birth/formation of stars/a star.	3%	0%	9%	0%	4%	5%	17%	0%	3%	3%	11%	0%
Student talks about the birth/formation of planets/a planet.	3%	0%	2%	0%	0%	0%	6%	0%	3%	0%	3%	0%
Student talks about the birth/formation of galaxies/a galaxy.	1%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
Student talks about the birth/formation of unspecified objects.	2%	0%	0%	0%	4%	0%	0%	0%	3%	0%	0%	0%
Student talks about the death of stars.	1%	5%	9%	0%	4%	16%	22%	0%	2%	10%	12%	0%
Student talks about the death of planets.	0%	0%	0%	0%	0%	0%	6%	0%	0%	0%	1%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of stars/a star.	13%	8%	4%	50%	21%	8%	0%	0%	15%	8%	3%	50%
Student talks about changes during the lives of planets/a planet.	69%	23%	23%	0%	46%	5%	6%	0%	63%	14%	19%	0%
Student talks about changes during the lives of galaxies/a galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of unspecified objects.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about how the universe is big.	0%	0%	12%	0%	7%	5%	11%	0%	2%	3%	12%	0%
Student talks about competing effects canceling out.	1%	3%	11%	0%	4%	0%	6%	0%	2%	1%	9%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	2%	18%	23%	0%	4%	11%	6%	0%	3%	14%	19%	0%
Student says s/he has no idea.	2%	5%	7%	0%	4%	3%	0%	0%	3%	4%	5%	0%
Response field left blank.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 7.35: Common reasoning elements used by non-lecture-tutorial students (pre-instruction) in their answers to Form C, Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	15%	67%	0%	0%	0%	67%	6%	0%	7%	67%	4%	0%
Student talks about the birth/formation of stars/a star.	46%	0%	29%	0%	21%	0%	28%	0%	33%	0%	28%	0%
Student talks about the birth/formation of planets/a planet.	15%	27%	14%	0%	0%	0%	0%	0%	7%	17%	4%	0%
Student talks about the birth/formation of galaxies/a galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of unspecified objects.	0%	0%	0%	0%	7%	0%	0%	0%	4%	0%	0%	0%
Student talks about the death of stars.	8%	7%	14%	0%	7%	0%	28%	0%	7%	4%	24%	0%
Student talks about the death of planets.	0%	7%	14%	0%	0%	0%	0%	0%	0%	4%	4%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	0%	0%	0%	0%	7%	0%	0%	0%	4%	0%	0%	0%
Student talks about changes during the lives of stars/a star.	15%	0%	14%	0%	7%	22%	6%	0%	11%	8%	8%	0%
Student talks about changes during the lives of planets/a planet.	23%	13%	14%	0%	50%	11%	17%	0%	37%	13%	16%	0%
Student talks about changes during the lives of galaxies/a galaxy.	0%	0%	0%	0%	7%	11%	0%	0%	4%	4%	0%	0%
Student talks about changes during the lives of unspecified objects.	15%	0%	0%	0%	7%	0%	0%	0%	11%	0%	0%	0%
Student talks about how the universe is big.	0%	0%	14%	0%	0%	0%	11%	0%	0%	0%	12%	0%
Student talks about competing effects canceling out.	0%	0%	29%	0%	0%	0%	28%	0%	0%	0%	28%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	8%	13%	29%	100%	7%	0%	6%	0%	7%	8%	12%	100%
Student says s/he has no idea.	8%	20%	0%	0%	0%	0%	0%	0%	4%	13%	0%	0%
Response field left blank.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 7.36: Common reasoning elements used by lecture-tutorial students (post-instruction) in their answers to Form C, Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	36%	87%	25%	0%	0%	87%	0%	0%	31%	87%	14%	0%
Student talks about the birth/formation of stars/a star.	0%	0%	0%	0%	50%	1%	0%	0%	6%	0%	0%	0%
Student talks about the birth/formation of planets/a planet.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of galaxies/a galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of unspecified objects.	0%	0%	0%	0%	0%	9%	0%	0%	0%	3%	0%	0%
Student talks about the death of stars.	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%
Student talks about the death of planets.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	0%	0%	0%	0%	0%	1%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of stars/a star.	0%	0%	0%	0%	0%	3%	67%	0%	0%	1%	29%	0%
Student talks about changes during the lives of planets/a planet.	36%	1%	0%	0%	0%	1%	0%	0%	31%	1%	0%	0%
Student talks about changes during the lives of galaxies/a galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of unspecified objects.	7%	0%	0%	0%	0%	0%	0%	0%	6%	0%	0%	0%
Student talks about how the universe is big.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about competing effects canceling out.	0%	0%	0%	0%	0%	0%	33%	0%	0%	0%	14%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	1%	25%	0%	0%	0%	0%	0%	0%	1%	14%	0%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Response field left blank.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 7.37: Common reasoning elements used by non-lecture-tutorial students (post-instruction) in their answers to Form C, Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	14%	50%	57%	0%	50%	68%	0%	0%	22%	64%	57%	0%
Student talks about the birth/formation of stars/a star.	43%	0%	57%	0%	0%	6%	0%	0%	33%	5%	57%	0%
Student talks about the birth/formation of planets/a planet.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of galaxies/a galaxy.	14%	0%	0%	0%	0%	0%	0%	0%	11%	0%	0%	0%
Student talks about the birth/formation of unspecified objects.	14%	0%	0%	0%	100%	0%	0%	0%	33%	0%	0%	0%
Student talks about the death of stars.	43%	13%	29%	0%	0%	0%	0%	0%	33%	2%	29%	0%
Student talks about the death of planets.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	14%	0%	0%	0%	0%	0%	0%	0%	11%	0%	0%	0%
Student talks about changes during the lives of stars/a star.	29%	0%	0%	0%	50%	0%	0%	0%	33%	0%	0%	0%
Student talks about changes during the lives of planets/a planet.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of galaxies/a galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of unspecified objects.	0%	0%	14%	0%	0%	0%	0%	0%	0%	0%	14%	0%
Student talks about how the universe is big.	0%	0%	29%	0%	0%	0%	0%	0%	0%	0%	29%	0%
Student talks about competing effects canceling out.	0%	0%	71%	0%	0%	0%	0%	0%	0%	0%	71%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	14%	0%	14%	0%	0%	6%	0%	0%	11%	5%	14%	0%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Response field left blank.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 7.38: Common reasoning elements used by students (pre-instruction) in their answers to Form C, Item 2: How does the total amount of matter in the universe right now compare to the total amount of matter in the universe at the very beginning of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says there was no matter/only energy in the beginning.	0%	1%	0%	0%	0%	0%
Student talks about the temperature cooling and/or that matter formed from energy.	1%	1%	1%	3%	0%	1%
Student says there is more matter now because the universe is expanding.	1%	5%	2%	0%	9%	5%
Student says the amount of matter increases as objects form and/or evolve.	23%	11%	19%	10%	19%	15%
Student says the amount of matter decreases as objects form and/or evolve.	7%	1%	5%	0%	7%	4%
Student says the amount of matter increases as objects interact and/or die.	2%	4%	3%	3%	5%	4%
Student says the amount of matter decreases as objects interact and/or die.	3%	2%	3%	3%	0%	1%
Student says the amount of matter does not change.	44%	66%	51%	72%	49%	60%
Student gives irrelevant information.	2%	2%	2%	0%	2%	1%
Student gives other reason not specified above.	7%	1%	5%	10%	2%	6%
Student says s/he has no idea.	2%	5%	3%	0%	5%	2%
Response field left blank or no reason given.	11%	8%	10%	5%	5%	5%

Table 7.39: Common reasoning elements used by students (post-instruction) in their answers to Form C, Item 2: How does the total amount of matter in the universe right now compare to the total amount of matter in the universe at the very beginning of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

Reasoning Element	Class D	Class E	LT Classes	Class F	Class G	Non-LT Classes
Student says there was no matter/only energy in the beginning.	53%	26%	44%	0%	3%	2%
Student talks about the temperature cooling and/or that matter formed from energy.	22%	17%	20%	0%	3%	2%
Student says there is more matter now because the universe is expanding.	6%	1%	5%	0%	14%	8%
Student says the amount of matter increases as objects form and/or evolve.	4%	13%	7%	5%	38%	25%
Student says the amount of matter decreases as objects form and/or evolve.	1%	1%	1%	0%	0%	0%
Student says the amount of matter increases as objects interact and/or die.	1%	4%	2%	0%	3%	2%
Student says the amount of matter decreases as objects interact and/or die.	0%	8%	3%	5%	3%	3%
Student says the amount of matter does not change.	18%	39%	25%	82%	27%	47%
Student gives irrelevant information.	1%	1%	1%	0%	0%	0%
Student gives other reason not specified above.	4%	0%	3%	5%	5%	5%
Student says s/he has no idea.	1%	0%	0%	0%	14%	8%
Response field left blank or no reason given.	9%	4%	7%	9%	14%	12%

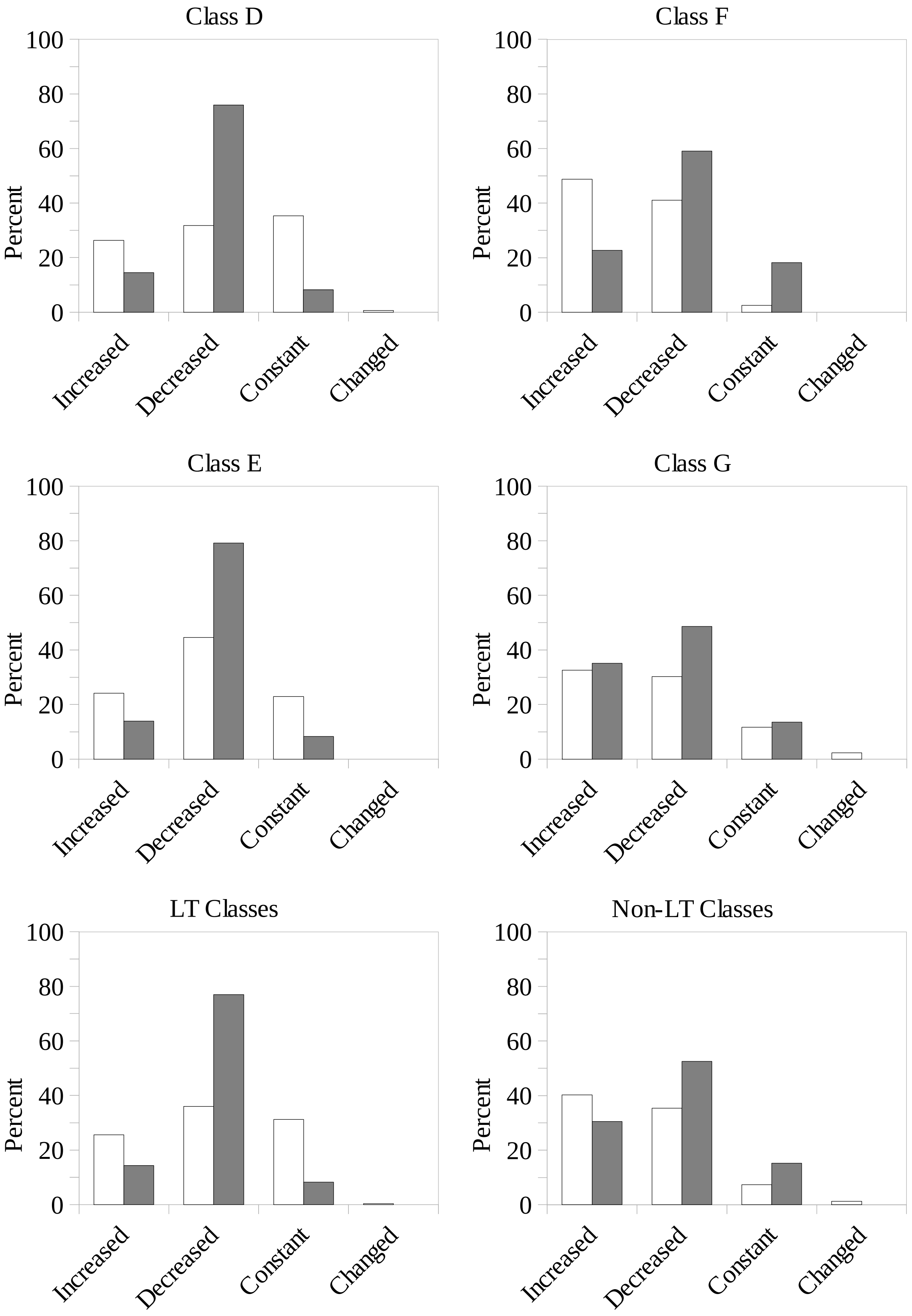


Figure 7.15: Percent of students who think the density of matter in the universe has increased, decreased, stayed constant, or changed (without specifying if it went up or down). White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Table 7.40: Common reasoning elements used by lecture-tutorial students (pre-instruction) in their answers to Form C, Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Students says the amount of matter in the universe has changed.	25%	0%	5%	0%	0%	0%	0%	0%	17%	0%	4%	0%
Student says the amount of matter in the universe has not changed.	0%	9%	22%	0%	0%	22%	11%	0%	0%	14%	19%	0%
Student says the size of the universe has changed/stuff spread out.	14%	36%	0%	0%	5%	84%	0%	0%	11%	56%	0%	0%
Student says the size of the universe has not changed/stuff hasn't spread out.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time.	16%	8%	2%	0%	35%	3%	0%	0%	22%	6%	1%	0%
Student talks about objects evolving over time.	27%	15%	2%	0%	25%	3%	0%	0%	27%	10%	1%	0%
Student talks about objects dying/breaking up.	0%	17%	5%	100%	5%	5%	0%	0%	2%	12%	4%	100%
Student talks about the effects of gravity.	2%	2%	2%	0%	0%	0%	0%	0%	2%	1%	1%	0%
Student talks about matter changing forms.	0%	4%	2%	0%	15%	0%	0%	0%	5%	2%	1%	0%
Student gives some other reason not specified above.	7%	6%	5%	0%	5%	3%	0%	0%	6%	4%	4%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says s/he has no idea.	9%	9%	19%	0%	10%	0%	11%	0%	9%	6%	17%	0%
Answer field is blank or no response given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 7.41: Common reasoning elements used by non-lecture-tutorial students (pre-instruction) in their answers to Form C, Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Students says the amount of matter in the universe has changed.	5%	0%	0%	0%	7%	8%	20%	0%	6%	3%	17%	0%
Student says the amount of matter in the universe has not changed.	0%	6%	100%	0%	0%	8%	20%	100%	0%	7%	33%	100%
Student says the size of the universe has changed/stuff spread out.	11%	75%	0%	0%	7%	77%	20%	0%	9%	76%	17%	0%
Student says the size of the universe has not changed/stuff hasn't spread out.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time.	47%	6%	0%	0%	64%	0%	0%	0%	55%	3%	0%	0%
Student talks about objects evolving over time.	21%	0%	0%	0%	14%	0%	0%	100%	18%	0%	0%	100%
Student talks about objects dying/breaking up.	5%	6%	0%	0%	0%	15%	0%	0%	3%	10%	0%	0%
Student talks about the effects of gravity.	16%	0%	0%	0%	7%	0%	0%	0%	12%	0%	0%	0%
Student talks about matter changing forms.	21%	6%	0%	0%	0%	0%	0%	0%	12%	3%	0%	0%
Student gives some other reason not specified above.	0%	0%	0%	0%	7%	0%	0%	0%	3%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says s/he has no idea.	0%	6%	0%	0%	0%	0%	0%	100%	0%	3%	0%	100%
Answer field is blank or no response given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 7.42: Common reasoning elements used by lecture-tutorial students (post-instruction) in their answers to Form C, Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Students says the amount of matter in the universe has changed.	74%	2%	8%	0%	20%	5%	17%	0%	58%	3%	11%	0%
Student says the amount of matter in the universe has not changed.	0%	0%	23%	0%	0%	12%	17%	0%	0%	4%	21%	0%
Student says the size of the universe has changed/stuff spread out.	13%	88%	31%	0%	20%	96%	17%	0%	15%	91%	26%	0%
Student says the size of the universe has not changed/stuff hasn't spread out.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time.	17%	0%	0%	0%	50%	0%	0%	0%	27%	0%	0%	0%
Student talks about objects evolving over time.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects dying/breaking up.	4%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%
Student talks about the effects of gravity.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about matter changing forms.	0%	0%	0%	0%	10%	2%	0%	0%	3%	1%	0%	0%
Student gives some other reason not specified above.	9%	1%	8%	0%	0%	2%	17%	0%	6%	1%	11%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Answer field is blank or no response given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Table 7.43: Common reasoning elements used by non-lecture-tutorial students (post-instruction) in their answers to Form C, Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	I	D	S	C	I	D	S	C	I	D	S	C
Students says the amount of matter in the universe has changed.	0%	8%	0%	0%	8%	0%	0%	0%	6%	3%	0%	0%
Student says the amount of matter in the universe has not changed.	0%	8%	0%	0%	0%	11%	0%	0%	0%	10%	0%	0%
Student says the size of the universe has changed/stuff spread out.	0%	69%	25%	0%	8%	89%	0%	0%	6%	81%	11%	0%
Student says the size of the universe has not changed/stuff hasn't spread out.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time.	80%	0%	50%	0%	54%	0%	0%	0%	61%	0%	22%	0%
Student talks about objects evolving over time.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects dying/breaking up.	0%	8%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%
Student talks about the effects of gravity.	20%	0%	0%	0%	0%	0%	0%	0%	6%	0%	0%	0%
Student talks about matter changing forms.	20%	0%	25%	0%	0%	0%	0%	0%	6%	0%	11%	0%
Student gives some other reason not specified above.	0%	0%	25%	0%	15%	0%	20%	0%	11%	0%	22%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Answer field is blank or no response given.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%

Items 4-6 on the spring 2010 version of Form C cover the same content as item 1 on its fall 2009 version. As with the items discussed above, students' graph choices (Figures 7.16-7.18) and responses (Tables 7.44-7.47 for item 4, Tables 7.48-7.51 for item 5, and Tables 7.52-7.55 for item 6) all reflect the same fundamental difficulty uncovered by the fall 2009 data: Some students struggle to incorporate the effects of expansion into their understandings of light-travel time and lookback time. The distribution of scores, graph choices, and responses on these items indicate that, in general, the lecture-tutorials help students with this difficulty. However, the data for item 6 suggests that students still have trouble thinking about lookback time in an expanding universe post-instruction. Despite the fact that the item tells students the age of the universe (13 billion years) and the time light has been traveling (8 billion years), many students fail to say that they are looking at a galaxy as it appeared when the universe was 5 billion years old. The data indicate that students cite the expansion of the universe as a reason not to choose the correct answer (B). Of course, Tables 7.52-7.55 also show that a large percent of students did not provide a reason for their answer, which opens the possibility that there may be other undetected factors influencing students' responses.

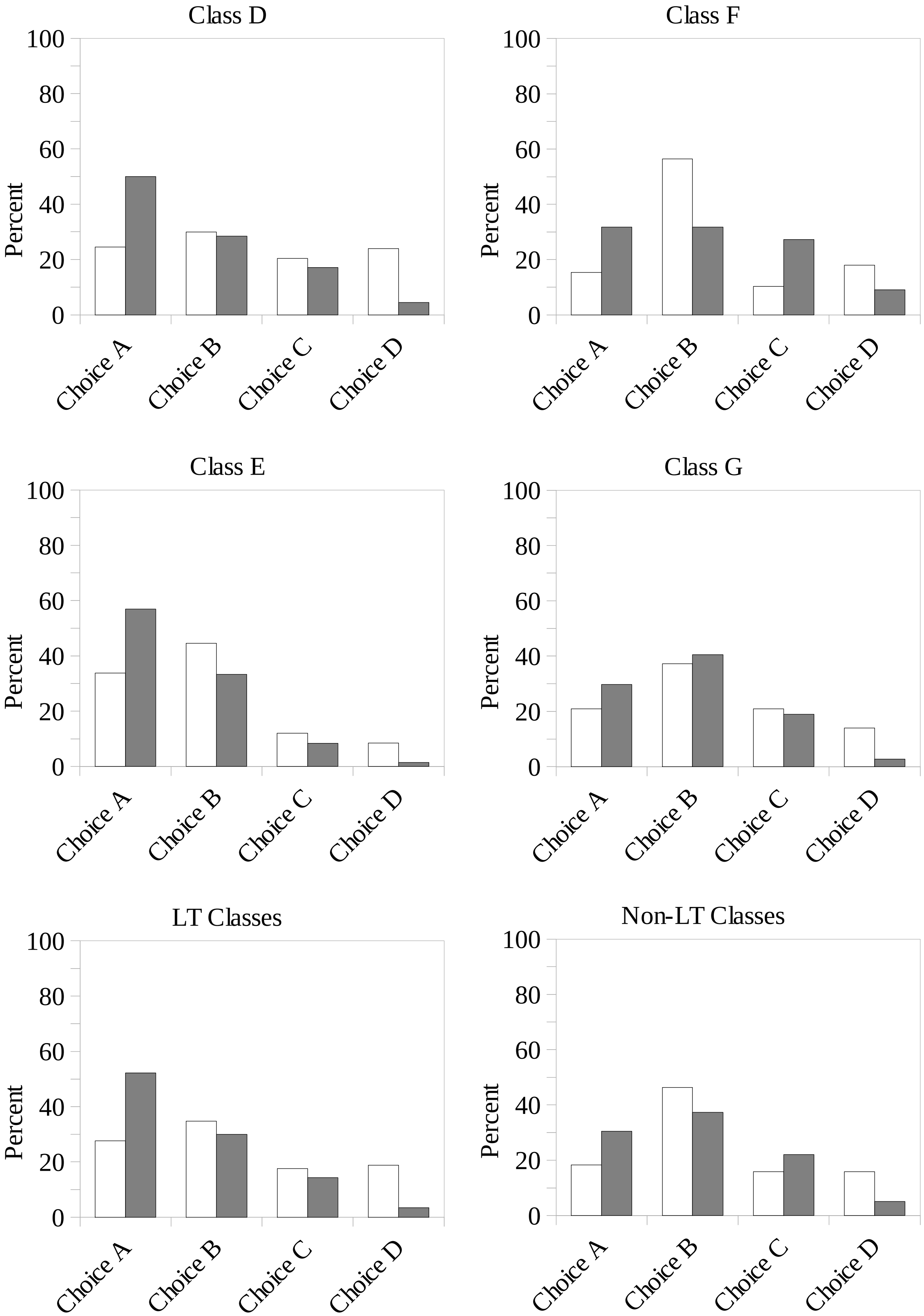


Figure 7.16: Percent of students who selected choice A, B, C, or D in response to item 4 on Form C. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Table 7.44: Common reasoning elements used by lecture-tutorial students (pre-instruction) in their answers to Form C, Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	54%	2%	18%	0%	68%	5%	30%	0%	59%	3%	20%	0%
Student talks about galaxies getting closer together.	2%	0%	3%	0%	0%	5%	0%	0%	1%	2%	2%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	2%	86%	3%	3%	0%	65%	0%	0%	1%	77%	2%	2%
Student says light-years are shorter than normal years.	5%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	3%	0%	0%	0%	10%	0%	0%	0%	5%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	2%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	2%	0%	3%	0%	7%	0%	0%	0%	4%	0%	2%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	5%	0%	0%	0%	43%	0%	0%	0%	11%
Student says we need more information on how galaxies moving together.	0%	2%	0%	3%	0%	0%	0%	29%	0%	1%	0%	6%
Student says 8 billion light-years is the minimum distance between Galaxies X and Y.	0%	0%	6%	0%	0%	0%	20%	0%	0%	0%	9%	0%
Student talks about time Galaxy X needs to see explosion.	29%	62%	38%	5%	36%	0%	30%	14%	32%	36%	36%	6%
Student says there's not enough information.	0%	0%	0%	88%	0%	0%	0%	43%	0%	0%	0%	81%
Student gives irrelevant information.	2%	0%	0%	3%	0%	0%	0%	0%	1%	0%	0%	2%
Student gives other reason not specified above.	12%	2%	47%	63%	14%	5%	0%	43%	13%	3%	36%	60%
Student says s/he has no idea.	0%	2%	0%	8%	0%	8%	10%	0%	0%	5%	2%	6%
Response field left blank or no reason given.	10%	8%	12%	5%	4%	19%	20%	0%	7%	13%	14%	4%

Table 7.45: Common reasoning elements used by non-lecture-tutorial students (pre-instruction) in their answers to Form C, Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	67%	5%	0%	0%	78%	0%	11%	0%	73%	3%	8%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	82%	0%	0%	0%	63%	0%	0%	0%	74%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	22%	17%	0%	0%	15%	8%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	43%	0%	0%	0%	33%	0%	0%	0%	38%
Student says we need more information on how galaxies moving together.	0%	0%	0%	43%	11%	13%	0%	17%	7%	5%	0%	31%
Student says 8 billion light-years is the minimum distance between Galaxies X and Y.	0%	0%	25%	0%	0%	0%	0%	0%	0%	0%	8%	0%
Student talks about time Galaxy X needs to see explosion.	83%	68%	0%	14%	22%	31%	0%	0%	47%	53%	0%	8%
Student says there's not enough information.	0%	0%	0%	86%	0%	0%	0%	67%	0%	0%	0%	77%
Student gives irrelevant information.	0%	0%	0%	14%	0%	0%	0%	0%	0%	0%	0%	8%
Student gives other reason not specified above.	33%	0%	50%	43%	11%	19%	22%	17%	20%	8%	31%	31%
Student says s/he has no idea.	0%	0%	0%	0%	0%	6%	0%	17%	0%	3%	0%	8%
Response field left blank or no reason given.	0%	18%	50%	0%	11%	13%	44%	17%	7%	16%	46%	8%

Table 7.46: Common reasoning elements used by lecture-tutorial students (post-instruction) in their answers to Form C, Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	87%	4%	41%	14%	93%	8%	17%	0%	89%	6%	36%	13%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	80%	7%	0%	0%	67%	0%	0%	0%	75%	6%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	17%	0%	0%	0%	3%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	13%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	100%	0%	0%	0%	13%
Student says 8 billion light-years is the minimum distance between Galaxies X and Y.	0%	0%	7%	0%	0%	0%	67%	0%	0%	0%	18%	0%
Student talks about time Galaxy X needs to see explosion.	56%	56%	41%	14%	34%	38%	50%	0%	48%	49%	42%	13%
Student says there's not enough information.	0%	0%	0%	57%	0%	0%	0%	0%	0%	0%	0%	50%
Student gives irrelevant information.	1%	0%	0%	0%	0%	0%	0%	0%	1%	0%	0%	0%
Student gives other reason not specified above.	5%	9%	26%	43%	2%	4%	0%	0%	4%	7%	21%	38%
Student says s/he has no idea.	1%	0%	11%	0%	0%	0%	0%	0%	1%	0%	9%	0%
Response field left blank or no reason given.	5%	11%	7%	43%	5%	25%	0%	0%	5%	16%	6%	38%

Table 7.47: Common reasoning elements used by non-lecture-tutorial students (post-instruction) in their answers to Form C, Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	100%	0%	50%	0%	73%	13%	29%	0%	83%	9%	38%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	57%	0%	0%	0%	80%	0%	0%	0%	73%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says 8 billion light-years is the minimum distance between Galaxies X and Y.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about time Galaxy X needs to see explosion.	71%	57%	17%	0%	36%	47%	29%	0%	50%	50%	23%	0%
Student says there's not enough information.	0%	0%	0%	50%	0%	0%	0%	0%	0%	0%	0%	33%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	14%	17%	0%	9%	0%	29%	0%	6%	5%	23%	0%
Student says s/he has no idea.	0%	0%	0%	0%	9%	0%	0%	0%	6%	0%	0%	0%
Response field left blank or no reason given.	0%	29%	50%	100%	9%	13%	29%	100%	6%	18%	38%	100%

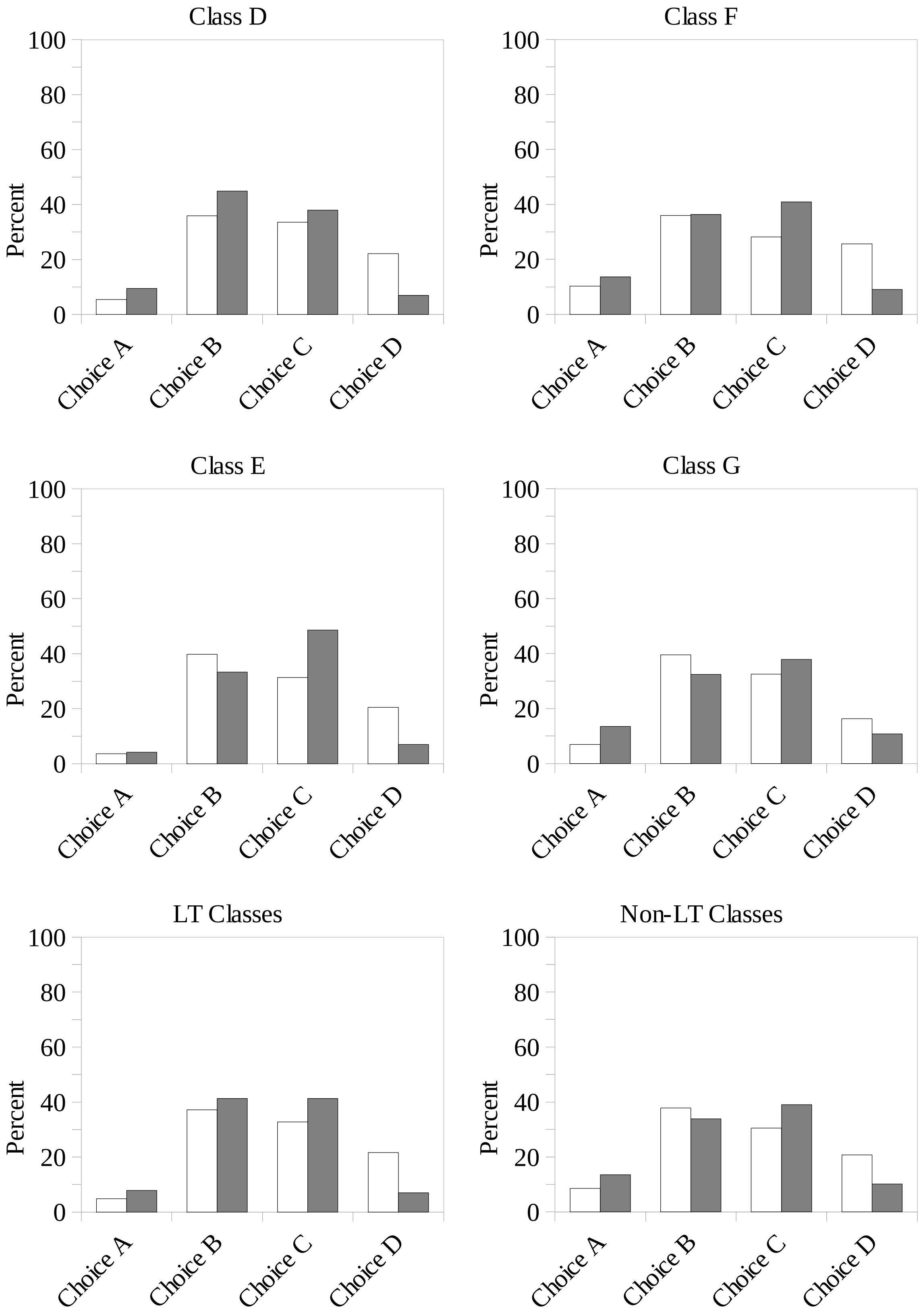


Figure 7.17: Percent of students who selected choice A, B, C, or D in response to item 5 on Form C. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Table 7.48: Common reasoning elements used by lecture-tutorial students (pre-instruction) in their answers to Form C, Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	33%	2%	45%	5%	0%	3%	62%	6%	25%	2%	50%	6%
Student talks about galaxies getting closer together.	0%	0%	2%	0%	33%	0%	0%	0%	8%	0%	1%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	75%	0%	0%	0%	58%	0%	0%	0%	69%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	2%	0%	0%	0%	4%	0%	0%	0%	2%	0%
Student says light-years are longer than normal years.	0%	0%	2%	0%	0%	0%	0%	0%	0%	0%	1%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	4%	0%	0%	0%	1%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	16%	0%	3%	0%	41%	0%	1%	0%	24%
Student says we need more information on how galaxies moving together.	0%	0%	0%	14%	0%	3%	0%	35%	0%	1%	0%	20%
Student says we need more information on the rate of expansion.	0%	0%	0%	5%	0%	0%	0%	0%	0%	0%	0%	4%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	0%	0%	3%	0%	0%	0%	6%	0%	0%	0%	4%
Student says there's not enough information.	0%	0%	0%	68%	0%	0%	0%	0%	0%	0%	0%	46%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	56%	3%	27%	43%	0%	0%	15%	6%	42%	2%	23%	31%
Student says s/he has no idea.	11%	5%	7%	16%	0%	6%	4%	6%	8%	5%	6%	13%
Response field left blank or no reason given.	0%	15%	18%	8%	67%	33%	12%	41%	17%	22%	16%	19%

Table 7.49: Common reasoning elements used by non-lecture-tutorial students (pre-instruction) in their answers to Form C, Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	25%	7%	73%	10%	33%	0%	43%	0%	29%	3%	56%	6%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	64%	0%	10%	0%	65%	0%	0%	0%	65%	0%	6%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	7%	0%	0%	0%	4%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	40%	0%	0%	0%	29%	0%	0%	0%	35%
Student says we need more information on how galaxies moving together.	0%	0%	0%	40%	0%	0%	0%	29%	0%	0%	0%	35%
Student says we need more information on the rate of expansion.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	50%	0%	20%	0%	0%	0%	0%	0%	23%	0%	12%
Student says there's not enough information.	0%	0%	0%	50%	0%	0%	0%	43%	0%	0%	0%	47%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	25%	0%	18%	10%	33%	6%	21%	0%	29%	3%	20%	6%
Student says s/he has no idea.	0%	0%	0%	10%	0%	6%	0%	43%	0%	3%	0%	24%
Response field left blank or no reason given.	50%	36%	9%	20%	33%	24%	29%	14%	43%	29%	20%	18%

Table 7.50: Common reasoning elements used by lecture-tutorial students (post-instruction) in their answers to Form C, Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	40%	0%	72%	0%	0%	4%	89%	0%	33%	1%	78%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	80%	0%	0%	0%	67%	0%	0%	0%	77%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	40%	0%	0%	0%	13%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	40%	0%	0%	0%	13%
Student says we need more information on the rate of expansion.	0%	0%	0%	0%	0%	0%	0%	40%	0%	0%	0%	13%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	0%	2%	9%	0%	4%	0%	20%	0%	1%	1%	13%
Student says there's not enough information.	0%	0%	0%	45%	0%	0%	0%	0%	0%	0%	0%	31%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	27%	4%	15%	45%	33%	8%	0%	0%	28%	5%	9%	31%
Student says s/he has no idea.	7%	3%	0%	18%	0%	4%	0%	0%	6%	3%	0%	13%
Response field left blank or no reason given.	27%	13%	12%	27%	67%	25%	11%	0%	33%	16%	12%	19%

Table 7.51: Common reasoning elements used by non-lecture-tutorial students (post-instruction) in their answers to Form C, Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	67%	0%	44%	0%	40%	0%	50%	0%	50%	0%	48%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	38%	0%	0%	0%	67%	0%	0%	0%	55%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	50%	0%	0%	0%	33%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	50%	0%	0%	0%	33%
Student says we need more information on the rate of expansion.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	38%	11%	0%	0%	0%	0%	25%	0%	15%	4%	17%
Student says there's not enough information.	0%	0%	0%	0%	0%	0%	0%	75%	0%	0%	0%	50%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	25%	0%	0%	0%	17%	14%	0%	0%	20%	9%	0%
Student says s/he has no idea.	0%	0%	0%	0%	40%	0%	0%	25%	25%	0%	0%	17%
Response field left blank or no reason given.	33%	50%	44%	100%	40%	17%	36%	0%	38%	30%	39%	33%

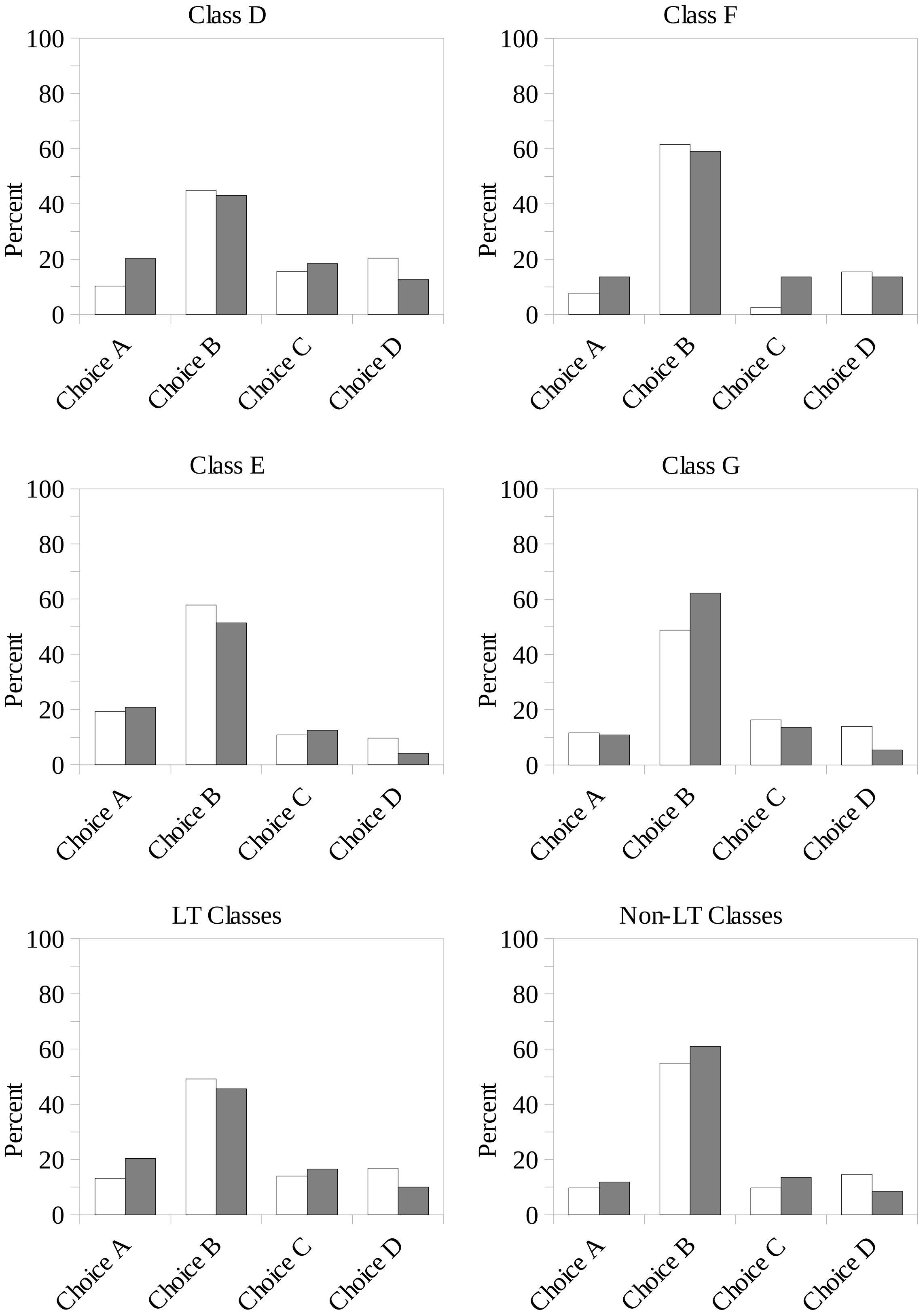


Figure 7.18: Percent of students who selected choice A, B, C, or D in response to item 6 on Form C. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Table 7.52: Common reasoning elements used by lecture-tutorial students (pre-instruction) in their answers to Form C, Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age? a) less than 5 billion years old; b) exactly 5 billion years old; c) more than 5 billion years old; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	6%	0%	4%	0%	13%	2%	11%	0%	9%	1%	6%	0%
Student talks about galaxies getting closer together.	6%	0%	0%	0%	0%	0%	0%	0%	3%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	12%	35%	8%	0%	0%	2%	0%	0%	6%	22%	6%	0%
Student says light-years are shorter than normal years.	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	3%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	4%	3%	0%	0%	0%	0%	0%	0%	3%	2%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	11%	0%	0%	0%	3%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	13%	0%	0%	0%	2%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	13%	0%	0%	0%	2%
Student reasons that 13-8=5/light traveled for 8 billion years.	24%	61%	8%	0%	19%	69%	11%	0%	21%	64%	9%	0%
Student talks about what Galaxy Y is like after the explosion.	0%	0%	0%	0%	0%	0%	11%	0%	0%	0%	3%	0%
Student talks about the star exploding more than 8 billion years ago.	6%	0%	4%	3%	6%	0%	0%	0%	6%	0%	3%	2%
Student talks about how it takes time to see events happen in the universe/lookback time.	12%	0%	0%	0%	13%	27%	11%	0%	12%	11%	3%	0%
Student says expansion affects age.	0%	0%	0%	0%	0%	0%	11%	0%	0%	0%	3%	0%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	50%	0%	0%	0%	13%	0%	0%	0%	43%
Student gives irrelevant information.	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	3%	0%
Student gives other reason not specified above.	29%	3%	19%	26%	6%	6%	0%	25%	18%	4%	14%	26%
Student says s/he has no idea.	6%	4%	8%	38%	6%	2%	11%	13%	6%	3%	9%	33%
Response field left blank or no reason given.	18%	17%	38%	12%	50%	19%	56%	38%	33%	18%	43%	17%

Table 7.53: Common reasoning elements used by non-lecture-tutorial students (pre-instruction) in their answers to Form C, Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age? a) less than 5 billion years old; b) exactly 5 billion years old; c) more than 5 billion years old; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	33%	0%	0%	0%	20%	0%	14%	0%	25%	0%	13%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	4%	0%	0%	0%	10%	0%	0%	0%	7%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	17%	0%	0%	0%	8%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	17%	0%	0%	0%	8%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student reasons that 13-8=5/light traveled for 8 billion years.	0%	67%	0%	0%	20%	57%	0%	0%	13%	62%	0%	0%
Student talks about what Galaxy Y is like after the explosion.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the star exploding more than 8 billion years ago.	33%	0%	0%	0%	0%	0%	0%	0%	13%	0%	0%	0%
Student talks about how it takes time to see events happen in the universe/lookback time.	0%	21%	0%	0%	0%	10%	14%	0%	0%	16%	13%	0%
Student says expansion affects age.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	0%	0%	17%	0%	0%	0%	0%	0%	0%	0%	8%
Student says there's not enough information.	0%	0%	0%	33%	0%	0%	0%	50%	0%	0%	0%	42%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	4%	100%	50%	20%	0%	43%	0%	13%	2%	50%	25%
Student says s/he has no idea.	0%	4%	0%	17%	0%	10%	0%	0%	0%	7%	0%	8%
Response field left blank or no reason given.	33%	17%	0%	33%	40%	29%	29%	50%	38%	22%	25%	42%

Table 7.54: Common reasoning elements used by lecture-tutorial students (post-instruction) in their answers to Form C, Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age? a) less than 5 billion years old; b) exactly 5 billion years old; c) more than 5 billion years old; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class D				Class E				LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	25%	0%	38%	0%	33%	0%	56%	0%	28%	0%	42%	0%
Student talks about galaxies getting closer together.	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	3%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	13%	35%	7%	0%	0%	3%	11%	0%	9%	24%	8%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student reasons that 13-8=5/light traveled for 8 billion years.	9%	41%	7%	0%	13%	76%	0%	0%	11%	53%	5%	0%
Student talks about what Galaxy Y is like after the explosion.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the star exploding more than 8 billion years ago.	6%	0%	0%	0%	33%	0%	33%	0%	15%	0%	8%	0%
Student talks about how it takes time to see events happen in the universe/lookback time.	6%	13%	3%	0%	13%	24%	22%	0%	9%	17%	8%	0%
Student says expansion affects age.	0%	0%	0%	0%	0%	0%	11%	0%	0%	0%	3%	0%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	0%	0%	0%	7%	0%	0%	0%	2%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	45%	0%	0%	0%	0%	0%	0%	0%	39%
Student gives irrelevant information.	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%	3%	0%
Student gives other reason not specified above.	28%	7%	14%	35%	13%	3%	11%	0%	23%	6%	13%	30%
Student says s/he has no idea.	13%	1%	7%	25%	0%	0%	0%	0%	9%	1%	5%	22%
Response field left blank or no reason given.	19%	26%	24%	30%	27%	14%	22%	100%	21%	22%	24%	39%

Table 7.55: Common reasoning elements used by non-lecture-tutorial students (post-instruction) in their answers to Form C, Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age? a) less than 5 billion years old; b) exactly 5 billion years old; c) more than 5 billion years old; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Class F				Class G				Non-LT Classes			
Reasoning Element	A	B	C	D	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	33%	0%	0%	0%	0%	0%	0%	0%	14%	0%	0%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	0%	0%	0%	0%	22%	0%	0%	0%	14%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student reasons that 13-8=5/light traveled for 8 billion years.	0%	46%	0%	0%	0%	30%	0%	0%	0%	36%	0%	0%
Student talks about what Galaxy Y is like after the explosion.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the star exploding more than 8 billion years ago.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about how it takes time to see events happen in the universe/lookback time.	0%	8%	0%	0%	0%	4%	20%	0%	0%	6%	13%	0%
Student says expansion affects age.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says answer depends on where exactly events happen in Galaxies X and Y.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	33%	0%	0%	0%	50%	0%	0%	0%	40%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	33%	0%	0%	33%	25%	9%	20%	50%	29%	6%	13%	40%
Student says s/he has no idea.	0%	0%	0%	33%	25%	4%	0%	0%	14%	3%	0%	20%
Response field left blank or no reason given.	67%	54%	100%	33%	50%	39%	60%	50%	57%	44%	75%	40%

7.4.4 Form D Responses

Tables 7.56 and 7.57 give the distribution of scores for the three items of Form D. Note that we excluded item 3 from our analysis since, as mentioned in previous chapters, it erroneously assumed that stars orbiting at the same speed must feel the same gravitational force. Tables 7.56 and 7.57 show that the lecture-tutorial students show greater improvements pre- to post-instruction than the non-lecture-tutorial students.

Table 7.56: Pre-instruction distribution of scores on Form D for spring 2010.

	Class D				Class E				LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	1%	95%	4%	-	0%	93%	7%	-	1%	95%	5%	-
Item 2	0%	34%	66%	-	0%	11%	89%	-	0%	30%	70%	-
Item 4	1%	38%	60%	1%	0%	33%	67%	0%	1%	38%	61%	1%
	Class F				Class G				Non-LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	0%	100%	0%	-	3%	76%	21%	-	1%	89%	10%	-
Item 2	0%	11%	89%	-	6%	15%	79%	-	3%	13%	85%	-
Item 4	5%	11%	84%	0%	9%	65%	26%	0%	7%	36%	57%	0%

Table 7.57: Post-instruction distribution of scores on Form D for spring 2010.

	Class D				Class E				LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	1%	38%	61%	-	0%	65%	35%	-	1%	41%	58%	-
Item 2	0%	23%	77%	-	0%	17%	83%	-	0%	22%	78%	-
Item 4	0%	35%	33%	32%	0%	26%	65%	9%	0%	34%	37%	29%
	Class F				Class G				Non-LT Classes			
	0	1	2	3	0	1	2	3	0	1	2	3
Item 1	5%	59%	36%	-	0%	76%	24%	-	2%	69%	29%	-
Item 2	0%	23%	77%	-	0%	32%	68%	-	0%	29%	71%	-
Item 4	0%	32%	41%	27%	0%	41%	30%	30%	0%	37%	34%	29%

Figure 7.19 shows the percent of students selecting each of the graph options for item 1. Just like the fall 2009 data, the spring 2010 data show that the lecture-tutorial students are much more likely to select the correct rotation curve for a spiral galaxy than the non-lecture-tutorial students. As noted in Chapter 6, this result is especially striking when one considers the fact that,

in principle, students can simply memorize this answer regardless of whether or not they actually understand anything about dark matter and rotation curves. However, students that have gone through the lecture-tutorial activity of interpreting rotation curves appear to be much more likely to remember the correct answer for this item. Furthermore, the lecture-tutorial students outperform the non-lecture-tutorial students on item 2, which asks them to use the rotation curve they chose in item 1 to rank the speeds of three stars. This data suggests that the lecture-tutorials are helping students.

Figure 7.20 shows the percent of students who select choices A, B, and C for item 4. The correct answer, choice C, is selected by the smallest percent of students in both populations pre-instruction. This suggests that many students think the matter in a spiral galaxy is predominantly located in its center or arms, which is a sensible conclusion if one does not know about dark matter. Post-instruction, the percent of students selecting choice C increases. A greater percentage of lecture-tutorial students select C post-instruction than non-lecture-tutorial students.

Item 4 also asks students to explain their reasoning for their choices. Tables 7.58 and 7.60 show the various reasoning elements used by students pre- and post-instruction, respectively. In general, students written responses were not very edifying. Many made references to gravity or to various parts of the galaxy. Post-instruction, many students also talked about dark matter. However, this item failed to elicit many responses that give us further insight into potential student difficulties. This was compounded by the fact that many students simply circled one of the answer choices without providing any written justification for their choices. Thus, while item 4 did provide us with some information, students' written responses marked this item as a candidate for revision.

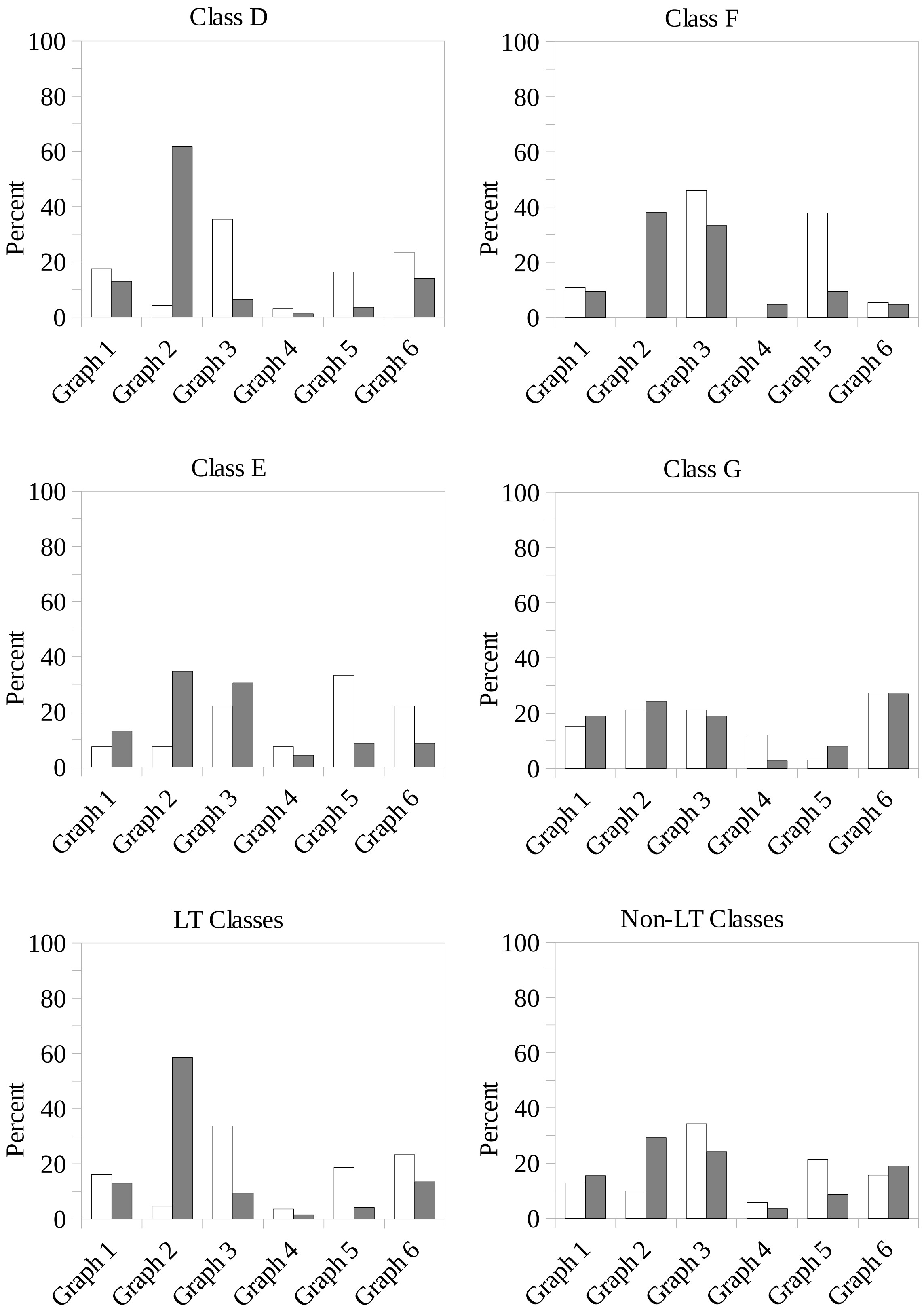


Figure 7.19: Students' graph choices for item 1 on Form D. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

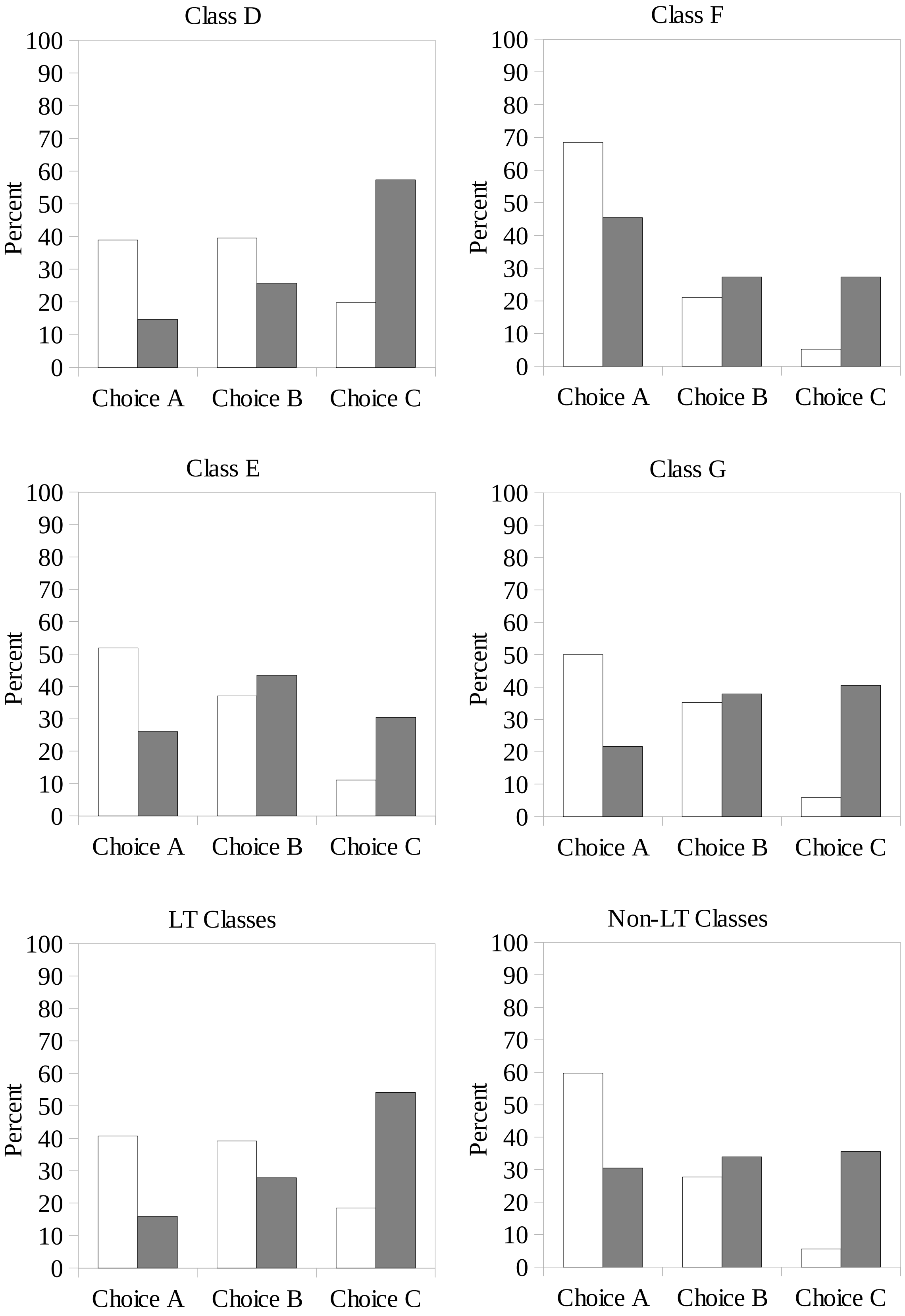


Figure 7.20: Percent of students who selected choice A, B, C, or D in response to item 4 on Form D. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Table 7.58: Common reasoning elements used by lecture-tutorial students (pre-instruction) in their answers to Form D, Item 4: Based on your previous answers, how is matter distributed in our galaxy? Pick the best answer from the following choices (a-c): a) Most of the matter in the galaxy is located in the center; b) Most of the matter in the galaxy is located in the center and spiral arms; c) Neither a nor b. Explain your reasoning for your choice.

	Class D			Class E			LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student talks about gravity.	38%	35%	6%	43%	30%	0%	48%	43%	7%
Student talks about unspecified forces or pulls.	5%	8%	0%	7%	0%	0%	6%	8%	0%
Student talks about spinning or revolving.	11%	6%	9%	21%	20%	0%	15%	10%	11%
Student talks about visible/normal matter.	2%	3%	0%	0%	0%	0%	2%	3%	0%
Student talks about unseen/dark matter.	0%	0%	3%	0%	0%	33%	0%	0%	7%
Student talks about the galaxy's central black hole.	2%	0%	3%	0%	10%	0%	2%	2%	4%
Student talks about stellar evolution.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about galactic evolution.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the center/bulge.	61%	57%	13%	50%	60%	33%	75%	72%	19%
Student talks about the spiral arms/disk.	3%	41%	13%	0%	40%	33%	3%	52%	19%
Student talks about the halo/what surrounds the galaxy.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about an even or equal distribution of matter.	0%	2%	67%	0%	0%	33%	0%	2%	78%
Student relates the location(s) of mass to observations or previous answers.	2%	2%	0%	21%	30%	0%	7%	7%	0%
Student talks about where the galaxy appears densest or brightest.	26%	19%	0%	7%	40%	0%	29%	27%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	13%	24%	17%	21%	10%	0%	19%	27%	19%
Student says s/he has no idea.	10%	3%	3%	0%	0%	0%	11%	4%	4%
Response field left blank.	10%	14%	14%	7%	0%	0%	13%	15%	15%

Table 7.59: Common reasoning elements used by non-tutorial students (pre-instruction) in their answers to Form D, Item 4: Based on your previous answers, how is matter distributed in our galaxy? Pick the best answer from the following choices (a-c): a) Most of the matter in the galaxy is located in the center; b) Most of the matter in the galaxy is located in the center and spiral arms; c) Neither a nor b. Explain your reasoning for your choice.

	Class F			Class G			Non-LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student talks about gravity.	35%	38%	0%	18%	0%	0%	28%	15%	0%
Student talks about unspecified forces or pulls.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about spinning or revolving.	0%	0%	0%	12%	0%	0%	5%	0%	0%
Student talks about visible/normal matter.	0%	0%	0%	29%	58%	50%	12%	35%	25%
Student talks about unseen/dark matter.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the galaxy's central black hole.	8%	0%	0%	6%	8%	0%	7%	5%	0%
Student talks about stellar evolution.	0%	0%	0%	12%	8%	50%	5%	5%	25%
Student talks about galactic evolution.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the center/bulge.	58%	50%	0%	47%	67%	50%	53%	60%	25%
Student talks about the spiral arms/disk.	8%	38%	0%	6%	58%	50%	7%	50%	25%
Student talks about the halo/what surrounds the galaxy.	0%	0%	0%	0%	8%	50%	0%	5%	25%
Student talks about an even or equal distribution of matter.	0%	0%	50%	0%	0%	50%	0%	0%	50%
Student relates the location(s) of mass to observations or previous answers.	4%	0%	0%	0%	0%	0%	2%	0%	0%
Student talks about where the galaxy appears densest or brightest.	4%	13%	0%	18%	17%	0%	9%	15%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	12%	13%	50%	6%	8%	0%	9%	10%	25%
Student says s/he has no idea.	4%	0%	0%	6%	0%	0%	5%	0%	0%
Response field left blank.	19%	13%	0%	12%	25%	0%	16%	20%	0%

Table 7.60: Common reasoning elements used by lecture-tutorial students (post-instruction) in their answers to Form D, Item 4: Based on your previous answers, how is matter distributed in our galaxy? Pick the best answer from the following choices (a-c): a) Most of the matter in the galaxy is located in the center; b) Most of the matter in the galaxy is located in the center and spiral arms; c) Neither a nor b. Explain your reasoning for your choice.

	Class D			Class E			LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student talks about gravity.	8%	0%	1%	33%	0%	0%	13%	0%	1%
Student talks about unspecified forces or pulls.	4%	2%	1%	0%	10%	0%	3%	4%	1%
Student talks about spinning or revolving.	4%	0%	1%	17%	0%	0%	6%	0%	1%
Student talks about visible/normal matter.	0%	5%	6%	17%	0%	29%	3%	4%	8%
Student talks about unseen/dark matter.	8%	7%	52%	0%	30%	71%	6%	11%	53%
Student talks about the galaxy's central black hole.	0%	0%	1%	0%	0%	0%	0%	0%	1%
Student talks about stellar evolution.	4%	9%	0%	0%	0%	0%	3%	7%	0%
Student talks about galactic evolution.	4%	2%	0%	0%	10%	0%	3%	4%	0%
Student talks about the center/bulge.	48%	50%	17%	67%	50%	14%	52%	50%	17%
Student talks about the spiral arms/disk.	8%	48%	12%	0%	50%	57%	6%	48%	15%
Student talks about the halo/what surrounds the galaxy.	0%	5%	9%	0%	0%	29%	0%	4%	10%
Student talks about an even or equal distribution of matter.	8%	9%	67%	0%	0%	29%	6%	7%	65%
Student relates the location(s) of mass to observations or previous answers.	0%	2%	16%	0%	10%	14%	0%	4 %	16%
Student talks about where the galaxy appears densest or brightest.	12%	11%	1%	0%	0%	0%	10%	9%	1%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	12%	11%	7%	17%	0%	0%	13%	9%	7%
Student says s/he has no idea.	0%	5%	1%	0%	0%	0%	0%	4%	1%
Response field left blank.	24%	9%	2%	17%	20%	0%	23%	11%	2%

Table 7.61: Common reasoning elements used by non-tutorial students (post-instruction) in their answers to Form D, Item 4: Based on your previous answers, how is matter distributed in our galaxy? Pick the best answer from the following choices (a-c): a) Most of the matter in the galaxy is located in the center; b) Most of the matter in the galaxy is located in the center and spiral arms; c) Neither a nor b. Explain your reasoning for your choice.

	Class F			Class G			Non-LT Classes		
Reasoning Element	A	B	C	A	B	C	A	B	C
Student talks about gravity.	20%	0%	0%	50%	7%	0%	33%	5%	0%
Student talks about unspecified forces or pulls.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about spinning or revolving.	0%	33%	0%	25%	0%	0%	11%	10%	0%
Student talks about visible/normal matter.	0%	0%	17%	13%	14%	7%	6%	10%	10%
Student talks about unseen/dark matter.	10%	0%	100%	13%	0%	73%	11%	0%	81%
Student talks about the galaxy's central black hole.	10%	0%	0%	13%	0%	0%	11%	0%	0%
Student talks about stellar evolution.	0%	0%	0%	0%	7%	0%	0%	5%	0%
Student talks about galactic evolution.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the center/bulge.	70%	83%	17%	63%	21%	7%	67%	40%	10%
Student talks about the spiral arms/disk.	0%	67%	17%	13%	29%	7%	6%	40%	10%
Student talks about the halo/what surrounds the galaxy.	0%	0%	50%	0%	0%	33%	0%	0%	38%
Student talks about an even or equal distribution of matter.	0%	17%	0%	0%	0%	13%	0%	5%	10%
Student relates the location(s) of mass to observations or previous answers.	0%	0%	17%	0%	0%	0%	0%	0%	5%
Student talks about where the galaxy appears densest or brightest.	0%	33%	0%	0%	0%	0%	0%	10%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	10%	17%	0%	25%	7%	7%	17%	10%	5%
Student says s/he has no idea.	0%	0%	0%	0%	0%	0%	0%	0%	0%
Response field left blank.	20%	0%	0%	0%	57%	7%	11%	40%	5%

7.5 Validity

In this section, we return to the question of whether or not our surveys are valid. This issue was initially addressed in Section 6.6. This section continues and builds upon the interpretive argument for the validity of our score interpretations in Section 6.6.

We will not readdress in detail every component of the interpretive argument delineated in Section 6.6. For some components, the data in this chapter do not add any new evidence. For example, our assumption that "each survey adequately covers the construct it is intended to measure" is predominantly supported by the design and expert review processes described in previous chapters. Other components, such as "students' responses can be reliably transformed into numerical scores" and "these scores can be used to find measurable differences between populations of students," are self-evidently supported by the data presented in this chapter. This section instead focuses on two issues: "Do Astro 101 students correctly read and interpret our survey items?" and "Are there any other factors about the courses we surveyed that could explain differences in students' responses, other than the lecture-tutorials?" The former issue was addressed in Chapter 6 but must be readdressed here since we revised several survey items before the spring 2010 semester. The second issue was not addressed in Chapter 6 since, at that time, we did not have a sufficient number of classroom observations to provide evidence one way or the other.

To help us determine whether or not students interpret our survey items as we intended, I conducted semi-structured, one-on-one, videotaped, think-aloud interviews (Otero and Harlow 2009) with six Astro 101 students during the spring 2010. There students are hereafter known by the pseudonyms Nina, John, Paul, Gayle, Calvin, and Brenda. I conducted these interviews following the same procedure outlined in Section 6.6 for the fall 2009 interviews. The only difference between the fall 2009 and spring 2010 interviews was that I used the spring 2010 versions of the cosmology surveys for the spring 2010 interviews. Table 7.62 shows which survey forms each student tackled as well as the order in which they were presented.

Overall, these six students had no issues interpreting most of our survey items as we intended.

Table 7.62: The forms each student responded to during his/her interview in the spring 2010. The numbers denote the order in which I presented the surveys to the student.

Student	Form A	Form B	Form C	Form D
Nina	1	3	-	2
John	3	-	1	2
Paul	-	1	2	3
Gayle	2	-	3	1
Calvin	3	2	1	-
Brenda	2	1	3	-

Three students (Nina, Gayle, and Brenda) did catch some minor typos on the surveys. Nina furthermore called attention to the fact that item 4 of Form B refers to “the BB” instead of “the Big Bang.” However, all the interviewed students read “BB” as “Big Bang” and an inspection of the written responses to this item indicates that the vast majority of students in this study also interpreted “BB” as “Big Bang,” so this does not appear to be a fatal flaw in the item.

That does not mean that item 4 on Form B was perfect. Brenda, Calvin, and John all found the phrase “surrounding the event” to be confusing. Brenda’s response to this item summarizes the difficulty experienced by all three of these students:

> “Surrounding the event? Like is this, I mean (sighs), this question is confusing. Surrounding the event, like, of the Big Ba, the Big Bang, the time when they thought it actually happened, like the moment of the Big Bang, or like right before or right after, like surrounding the event? Or s- like, I don’t know what this question’s asking. I mean I guess I could assume that it means surrounding like the moment of the Big Bang or like the moment prior, like immediately prior to would be like a region of space that includes nothing because everything that was in the universe was like in this one little space and then expanded. I don’t know, that question’s super confusing.”

The fact that this same issue was raised by multiple students highlights this item as a candidate for revision.

There was another item that confused multiple students. Nina, John, Gayle, Calvin, and Brenda all misinterpreted item 6 on Form A. The problematic part of this item is the term “older universe.” Nina, John, Gayle, and Brenda all thought it meant what this universe will be like in the future. Calvin thought it referred to what the universe was like in the past. This confusion marked this item as a candidate for revision.

To help us determine whether or not there are factors other than instruction (with and without the lecture-tutorials) that might explain differences in students' performances on our surveys, I observed almost every lecture period of Classes E (which used the cosmology lecture-tutorials) and F (which did not use the cosmology lecture-tutorials). Specifically, I observed 95% of the Class E's lectures and 92% of Class F's lectures. During each lecture, I took detailed notes on the topics covered, the instructional methods used (e.g. lecture, lecture-tutorials, think-pair-share questions, interactive demonstrations, etc.), and the time spent on each topic and method. Following standard procedure for qualitative research (Erickson 1986; Patton 1980), I wrote down my reflections of each lecture immediately after each class or as soon after a class as possible.

Did the Class E, which used the lecture-tutorials, spend more time on cosmology than Class F, which did not? According to my records, Class E spent 327 minutes out of the 1751 minutes that I observed covering cosmology. This was spread out over nine lectures and accounts for 19% of the time I observed Class E. 29% of Class E's time on cosmology was spent doing the lecture-tutorials. In comparison, Class F spent 286 minutes out of the 1450 I observed on cosmology. This means that 20% of the time I observed Class F was devoted to cosmology. This cosmology material was spread out over 8 days, one of which I missed due to illness. Including this missed class does not dramatically effect my estimates of the percent of time Class F spent on cosmology. If I assume that 0 minutes of the approximately 90 minute lecture I missed was devoted to cosmology (which was not the case), then the percent of time Class F spent on cosmology changes to 19%. If I assume all 90 minutes were devoted to cosmology (a more likely scenario), then the percent changes to 24%. In any case, I have no evidence that Class E spent more class time overall on cosmology than Class F, despite the fact that Class E also incorporated the cosmology lecture-tutorials.

Class E did differ from Class F in that the instructor of Class E used more interactive engagement techniques throughout the course. 37% of the time I observed Class E was spent on interactive engagement. This included the cosmology lecture-tutorials, the original *Lecture-Tutorials for Introductory Astronomy*, think-pair-share questions, and interactive demonstrations.

In contrast, only 8% of the time I observed Class F was spent on interactive engagement, all of which involved think-pair-share questions. We do not have sufficient data to disentangle the effect these other interactive engagement techniques had on students' performances compared to the effect of the cosmology lecture-tutorials themselves.

However, I did videotape multiple groups working on the cosmology lecture-tutorials in Class E. I made these recordings in order to check whether or not the lecture-tutorials were prompting the kinds of discussions and promoting the learning of the kinds of topics we intended. In general, they were. For example, consider the exchange between two students (given the pseudonyms Will and Jose) answering the final question of the "Dark Matter" lecture-tutorial:

> Will: So, according to their, like, previous, what they previously thought, if they saw all these things traveling at the same speed, they would have thought they were at the same distance, but they're not, so ((inaudible)).
>
> Jose: Right. Well, they thought that the farther, the farther you got away from the mass -
>
> Will: The slower the speed.
>
> Jose: - the slower you would be orbiting, but it turns out what they found is that they were orbiting at the same speed.
>
> Will: Okay.
>
> Jose: So that means there's, it's either more evenly distributed from the center or there's, or there's, it's, it's like evenly distributed throughout or there's like the mass that we can't see that's messing with the orbit. So there's more mass in the halo then we can see. So we're assuming that all the mass that we're seeing, we're assuming that that's more massive because it's producing more light, but that's not necessarily taking into account the mass that might not be giving off light.

This dialog is typical of the conversations we observe from students working on the lecture-tutorials.

We cannot claim, however, that the lecture-tutorials always functioned as we intended. There was one case in which students finished a lecture-tutorial with an incorrect conclusion. At the end of the "Hubble's Law" lecture-tutorial, two students (pseudonyms Jane and Alexis) had the following conversation:

> Jane: Based on the Hubble plot in figure five, is the expansion rate of the universe increasing or decreasing as time goes on?
>
> Alexis: I think it's -
>
> Jane: Decreasing -
>
> Alexis: - decreasing.
>
> Jane: - I guess, 'cause eventually, I think, everything's going to get to that distance, that distance from us.
>
> Alexis: mm-hmm
>
> Jane: And so they're all going to slow down.
>
> Alexis: Yeah. Okay.
>
> Jane: It's not flatter in the past and steeper now.

These students completed the lecture-tutorial and arrived at the incorrect conclusion that the Hubble plot for our universe shows that the expansion rate of the universe has decreased over time. The student debate at the end of the lecture-tutorial did not make them change their answer. If this is typical of conversations on this lecture-tutorial, then it may explain why even the lecture-tutorial students do not show much gain on items 3 and 4 of Form A.

As one final check on our lecture-tutorials, I also showed some of the statements from the lecture-tutorials' students debates to the students I interviewed. Each time one of the interviewed students answered an item with a response similar to one of the statements from a student debate, I handed the statement to the interviewed student and asked him/her whether or not it accords with his/her opinion. The interviewed students agreed in all cases. This supports the idea that our student debates accurately reflect what real students think and say.

Overall, the data we collected in the spring 2010 adds to the evidence in support of our validity claims. Most of the survey items passed the review of the interviewed students. We could not find any difference in the time spent on cosmology between one of the lecture-tutorial classes and one of the non-lecture-tutorial classes. With one exception, the lecture-tutorials appear well-matched to the abilities of Astro 101 students and foster the kinds of discussions and learning we intended. While we cannot absolutely prove every assumption in our interpretive argument (as noted in Chapter 6), these findings give us greater confidence in the validity of our score interpretations.

Of course, the data also indicate that our surveys and lecture-tutorials were not perfected by the spring of 2010. Item 6 of Form A and item 4 of Form B were confused and misinterpreted by students. Videotape of students working on the "Hubble's law" lecture-tutorial demonstrated that students can easily complete the tutorial with some incorrect ideas intact. These findings prompted us to make further revisions to our surveys and lecture-tutorials in preparation for the fall 2010 semester.

7.6 Revisions

This section describes the revisions we made to both the surveys and the lecture-tutorials between the spring 2010 and fall 2010 semesters. Although both improved after the revisions we made before the spring 2010 semester, the data described above pointed out several areas that required further work. Section 7.6.1 describes the revisions we made to our surveys while Section 7.6.2 focuses on the lecture-tutorials.

7.6.1 Surveys

We made minor changes to all of the surveys in order to corrected typos and improve their clarity. A few items, however, received more substantial revisions and we completely overhauled Form D.

Item 6 on Form A and items 4 and 7 on Form B were all rewritten. Since the term "older universe" was problematic for many of the interviewed students, we removed that term entirely from the new version of item 6 on Form A. This new version now reads as follows:

> Use the blank graph provided below to draw what you think Figure 1 would look like for our universe if it took much longer to reach its current size. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

We also completely changed item 4 on Form B in response to student confusion with the phrase "surrounding the event." We changed this item to the following question:

> Circle the sentence that best describes the universe at the time of the Big Bang:
>
> a) In the beginning, there was space in the universe surrounding the location of the Big Bang but this space was empty of all matter.
>
> b) In the beginning, there was space in the universe surrounding the location of the Big Bang and matter already existed in this space.
>
> c) I think of the Big Bang differently than a or b.
>
> Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

Finally, we modified item 7 of Form in response to Brenda, who found the format somewhat awkward, and Calvin, who wanted more than just the three original options. Item 7 now reads as follows:

> Which of the following statements (a - d) are true? Circle all that apply. In general, the expansion of the universe causes ____.
>
> a) the distances between planets in the solar system to increase.
>
> b) the distances between stars in the galaxy to increase.
>
> c) the distances between galaxies in the universe to increase.
>
> d) None of the above.
>
> Explain your reasoning for your choice(s).

No other item on Forms A-C received substantial modifications.

Form D, however, was completely altered. We overhauled Form D for two reasons. First, by the end of the spring 2010 semester, we realized our error in assuming that stars with similar orbital speeds feel similar net gravitational forces. Second, we felt that we needed additional items to probe students' understandings of dark matter and rotation curves, especially in light of the low values of Cronbach's α reported in this and the previous chapter. These factors combined prompted us to change Form D into a new seven item survey.

Form D still begins with the same selection of rotation curves, but the first two items now ask students to pick which rotation curve corresponds to a solar system and which corresponds to a spiral galaxy. The next two items ask students to rank the relative orbital speeds of planets in a solar system and stars in a galaxy, respectively, based on these rotations curves. Two more items ask students where the majority of mass in a solar system and a spiral galaxy, respectively, are located. The final item on the new Form D asks students to compare planets orbiting a star and stars orbiting a galaxy and state whether or not the two situations are similar. We hoped that these new items would provide us additional insights into students ideas about rotation curves and dark matter.

The revisions described above constitute the changes we made to our surveys prior to the fall 2010 semester. Copies of the fall 2010 surveys are found in Appendix G. The scoring rubrics for these surveys are in Appendix I.

7.6.2 Lecture-Tutorials

In addition to the surveys, we also made several changes to the lecture-tutorials. As with our last set of revisions described in Chapter 6, most of these revisions involved minor tweaks to the wording and presentation of questions in order to improve their clarity. For example, we eliminated from the "Making Sense of the Universe and Expansion" lecture-tutorial question 5 and its accompanying figure since we felt it was not addressing the central point of the lecture-tutorial. We also revised question 13 on the same lecture-tutorial to clarify that the listed properties are properties of the real universe and to remove property d, which was just confusing many students. On the "Expansion, Lookback Times, and Distances" lecture-tutorial we gave a more explicit definition of lookback time. We also cleaned up many of the figures and tables in the lecture-tutorials. These are a few of the many smaller changes we made to the lecture-tutorials prior to the fall 2010.

We also modified two of the lecture-tutorials in order to improve their scientific accuracies. On the "Big Bang" lecture-tutorial we removed all references to the universe once being "small," since such references convey the idea that the universe once had a finite size. For some questions, we instead talk about regions of the universe once being small. We also reworked the "Dark Matter" lecture-tutorial to remove all questions that imply that objects with the same orbital speed feel the same net gravitational force. We intend these revisions to improve the content accuracy of the cosmology lecture-tutorials.

We also added a student debate to the "Dark Matter" lecture-tutorial. We added this debate after watching the videotape of students working on this lecture-tutorial in class. The students appeared to get tripped up on the very first question, which asks where most of the matter in the Solar System is located. The students wanted to say "in the Sun and the planets" instead of just "in the Sun." The new student debate is designed to call students' attentions to the fact that the planets make insignificant contributions to the total mass of the Solar System compared to the Sun.

Finally, we added several new questions to the "Hubble's Law" lecture-tutorial. Despite the fact that we revised this lecture-tutorial after the fall 2009 semester, lecture-tutorial students in the spring of 2010 still exhibited low gains and post-instruction performances on many of Form A's items on interpreting Hubble plots, as discussed above. Furthermore, our videotape of Jane and Alexis working on this lecture-tutorial revealed that it is possible for students to complete the lecture-tutorial with an incorrect idea intact. These observations lead us to include more questions designed to address students' understandings of Hubble plots.

The first of these new questions asks students to draw a Hubble plot for a universe whose expansion rate is zero. The next question asks students to draw a Hubble plot for a universe expanding at a faster rate over time. We expect students to draw the wrong Hubble plot for this questions. In an example of elicit-confront-resolve (Heron 2004b; McDermott 1991) in action, we then present students with the Hubble plot for our universe (which is Figure 7 in the lecture-tutorial) and ask them the following series of questions:

> 15) On Figure 7, draw a circle around the galaxies from which we receive information closest to our present time.

> 16) On Figure 7, draw a square around the galaxies from which we receive information furthest from our present time.
>
> 17) On Figure 7, write the letter A by the galaxies that are moving away from us with the fastest velocities.
>
> 18) On Figure 7, write the letter B by the galaxies that are moving away from us with the slowest velocities.
>
> 19) On Figure 7, write the letter S where the graph has the steepest slope.
>
> 20) On Figure 7, write the letter F where the graph has the flattest slope.
>
> 21) On Figure 7, write the Greek letter α by the portion of the graph that corresponds with the fastest expansion rate.
>
> 22) On Figure 7, write the Greek letter β by the portion of the graph that corresponds with the slowest expansion rate.
>
> 23) Based on the Hubble plot shown in Figure 7, would you say that the expansion rate of the universe is constant or changing with time? Explain your reasoning.
>
> 24) Based on the Hubble plot in Figure 7, is the expansion rate represented by the motion of galaxies far away from us faster than, slower than, or the same as the expansion rate represented by the motions of nearby galaxies? Explain your reasoning.
>
> 25) Based on the Hubble plot in Figure 7, is the expansion rate of the universe increasing or decreasing as time goes on? Explain your reasoning.
>
> [After a student debate, students then must answer the following question.]
>
> 27) Based upon your previous answers, is the graph you drew in Question 14 correct or does it need to be redrawn? Explain your reasoning.

We hoped the above questions would improve students' abilities to interpret Hubble plots.

Copies of these revised versions of the lecture-tutorials are located in Appendix H. These were the versions used by some of the participating students during the fall 2010 semester.

7.7 Summary of Spring 2010 Results

This chapter presented our results from the spring 2010 semester. In general, the spring 2010 data are consistent with our results from the fall 2009, which are detailed in Chapter 6. The students in the spring 2010 data set, in general, experienced the same difficulties as the fall 2009. We also saw similar patterns with regards to the efficacy of the lecture-tutorials: While CTT and IRT analyses show that students who use the lecture-tutorials outperform their peers who do not on Forms B-D, we can detect no effect on student performance on Form A. We cannot claim that students did any better on Form A if they used the lecture-tutorials. This is despite the fact that we revised the “Hubble’s Law” lecture-tutorial prior to the spring 2010 semester in order to address students’ difficulties on Form A.

We also collected additional data to support the validity of our interpretations of survey scores. Our six interviewed students, by in large, interpreted the spring 2010 survey items as we intended. We also used our observations of two classes (one of which used the lecture-tutorials and one of which did not) to argue against the possibility that the greater gains of lecture-tutorial students can be attributed to more time on cosmology than their non-lecture-tutorial counterparts. Finally, we discussed how our videotapes of students working on the lecture-tutorials in class reveals that the lecture-tutorials appear to be functioning as we intended. Taken together, the data we have collected seem to support the hypothesis that using the lecture-tutorials can lead to greater learning gains for Astro 101 students.

Our interviews and recordings of students working on the lecture-tutorials, combined with our more quantitative analyses, did reveal the need for additional revisions to both our surveys and the lecture-tutorials. We described the revisions we made based on the data in Section 7.6. We made these revisions prior to the fall 2010 semester, which is when we began collecting data from multiple institutions of various types across the United States. The following chapter presents the results from this fall 2010 data.

Chapter 8

Fall 2010 Results

The fall of 2010 was the final semester for which we collected data for this study. We collected data from fourteen different classes spread across eleven colleges and universities. These colleges and universities cover a range of higher education institution types, from community colleges to liberal arts colleges to large state-supported research universities. All participating students from this semester responded to the revised surveys described at the end of Chapter 7 (see Appendix G for the surveys and Appendix I for their scoring rubrics) and a subset used the revised versions of the cosmology lecture-tutorials. This chapter presents the results from the fall 2010 semester.

As with Chapter 7, I will not dwell on details discussed in previous chapters. I instead focus on where the fall 2010 data offer new insights into student difficulties and the efficacy of the cosmology lecture-tutorials.

This chapter is divided into six sections. Section 8.1 describes the participating classes. Section 8.2 gives our CTT analysis of the data. Our IRT analysis is covered in Section 8.3. Section 8.4 gives the details on students' responses to individual survey questions. Section 8.5 continues the validity argument begun in previous chapters. Section 8.6 summarizes this chapter's results.

8.1 Surveyed Classes

Table 8.1 shows the number of students who completed survey forms pre- and post-instruction in each of the fourteen classes in the fall of 2010. Table 8.1 also shows which classes used the lecture-tutorials and which did not. A total of 602 students pre-instruction and 554 students

Table 8.1: Number of participants pre- and post-instruction per class for spring 2010.

Class	Pre-Instruction	Post-Instruction	Used Lecture-Tutorials?	Survey Forms
Class H	110	92	Yes	A, B, C, D
Class I	52	67	Yes	A, B, C, D
Class J	20	20	Yes	C
Class K	21	16	Yes	D
Class L	8	4	Yes	C
Class M	34	30	Yes	B, D
Class N	9	7	Yes	B
Class O	10	8	Yes	C
Class P	120	106	Yes	A, B, D
Class Q	80	65	No	A, B, C, D
Class R	57	66	No	A, B, C, D
Class S	6	4	No	C
Class T	57	53	No	B, C, D
Class U	18	16	No	A

post-instruction took our surveys.

Nine of the fourteen classes used the cosmology lecture-tutorials. Two of these classes were large-enrollment Astro 101 courses taught at large state-supported research universities: Class P (at the University of California at Davis) and Class H (at the University of Colorado at Boulder). Class H was taught by the same instructor as Class B in the fall 2009; the only difference between Class H and Class B was that Class H lacked recitation sections. Two more classes were also taught at state-supported public universities: Class N (at Wichita State University) and Class O (at Towson University). One class (Class M) was taught at a public liberal arts institution (Washburn University). The rest (Class I, Class J, Class K, and Class L) were all taught at junior or community colleges (Santa Rosa Junior College, MiraCosta Community College, Suffolk County Community College, and Truckee Meadows Community College, respectively). We received a total of 384 surveys pre-instruction and 350 surveys post-instruction from students enrolled in these lecture-tutorial classes.

The remaining five classes did not use the lecture-tutorials. Two of these were taught at state universities: Class Q at the University of Colorado at Boulder and Class T at the University of Wisconsin at Parkside. Class U was taught at a private liberal arts college (Albion College). The other two classes were taught at community colleges: Class R at Santa Rosa Junior College (by the same instructor as Class I) and Class S at Truckee Meadows Community College. Overall, we received 218 surveys pre-instruction and 204 surveys post-instruction from these classes.

Table 8.1 shows which survey forms were administered in each participating class. Some classes were large enough that all four forms were administered, as was the case for classes in previous semesters. Other classes were small enough, however, that we choose just a subset of the forms to give to their students. We also decided which forms to administer in a given class based on the lecture-tutorials the instructor used. We always used the same form for a class post-instruction as we did pre-instruction.

Table 8.1 shows that the number of participating students in some of these classes is very small. Because of these small numbers, we hesitate to analyze these classes individually since such analyses may be plagued by the uncertainties surrounding small number statistics. Furthermore, presenting statistics for each class individually may be overwhelming and divert attention from our mission of comparing the lecture-tutorial classes to the non-lecture-tutorial classes *en masse*. We also feel that the data presented in the previous two chapters provides readers with a sufficient idea of how individual classes can vary in their responses to our surveys. For these reasons, we restrict our analyses and presentation of the data in Sections 8.2-8.4 to aggregate results for the lecture-tutorial classes and non-lecture-tutorial classes.

8.2 Classical Test Theory Analysis

As in previous chapters, our analysis begins by applying classical test theory to our data. Section 8.2.1 looks at the items' difficulties and discriminations. Section 8.2.2 examines the reliability of our surveys from a CTT perspective. Section 8.2.3 presents the normalized gains for each item and for each of the four surveys overall. Readers who need an overview of classical test theory are again referred back to Section 6.3.1.

Table 8.2: The discriminations of the items on Forms A-D for fall 2010.

Form A		Form B		Form C		Form D	
Item	Discrimin.	Item	Discrimin.	Item	Discrimin.	Item	Discrimin.
Item 1	0.53	Item 1	0.59	Item 1	0.62	Item 1	0.60
Item 2	0.45	Item 2	0.53	Item 2	0.35	Item 2	0.59
Item 3	0.38	Item 3	0.48	Item 3	0.69	Item 3	0.68
Item 4	0.39	Item 4	0.47	Item 4	0.51	Item 4	0.67
Item 5	0.69	Item 5	0.58	Item 5	0.55	Item 5	0.72
Item 6	0.73	Item 6	0.66	Item 6	0.46	Item 6	0.63
		Item 7	0.48			Item 7	0.72

8.2.1 Items' Difficulties and Discriminations

Table 8.2 shows the discriminations of the items on Forms A-D. In general, the discriminations on Forms A-C are approximately equal to the discriminations calculated for the analogous items for the spring 2010. A few items, however, have lower values, such as the first four items of Form A, item 6 of Form B, and items 2 and 6 of Form C. Nevertheless, the discriminations for all items fall within conventionally accepted bounds and thus present no obvious cause for concern.

Table 8.3 shows the items' P-values. The overall P-values for the items all fall within conventionally accepted limits. Examining the pre- and post-instruction P-values for the lecture-tutorial and non-lecture-tutorial populations reveals three interesting patterns.

First, the differences in the post- and pre-instruction P-values on Form D's items for the lecture-tutorial population are larger than the differences for the non-lecture-tutorial population. This pattern holds for every item on Form D. This suggests that students who use the cosmology lecture-tutorials outperform their peers who do not on this revised and expanded version of Form D.

A similar case can be made for some of the items on Form A. The change in the P-values pre- to post-instruction is greater for the lecture-tutorial population on items 2-5 than for the non-lecture-tutorial population. This is consistent with the hypothesis that the lecture-tutorials improve students' performances on Form A more than lecture alone. However, the differences here are smaller than on Form D. This, coupled with the fact that both the lecture-tutorial and non-

Table 8.3: The difficulties (P-values) of the items on Forms A-D for fall 2010.

	Item	Overall	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Form A	Item 1	0.47	0.46	0.45	0.54	0.48
	Item 2	0.39	0.38	0.41	0.41	0.40
	Item 3	0.35	0.34	0.42	0.33	0.33
	Item 4	0.36	0.34	0.44	0.33	0.33
	Item 5	0.81	0.80	0.85	0.79	0.80
	Item 6	0.64	0.64	0.65	0.56	0.67
Form B	Item 1	0.57	0.48	0.71	0.48	0.65
	Item 2	0.65	0.59	0.72	0.57	0.71
	Item 3	0.80	0.72	0.92	0.75	0.90
	Item 4	0.43	0.38	0.49	0.36	0.55
	Item 5	0.48	0.41	0.60	0.35	0.61
	Item 6	0.51	0.44	0.59	0.37	0.68
	Item 7	0.75	0.72	0.83	0.68	0.83
Form C	Item 1	0.64	0.51	0.77	0.45	0.76
	Item 2	0.46	0.42	0.54	0.42	0.48
	Item 3	0.57	0.46	0.64	0.44	0.69
	Item 4	0.52	0.49	0.59	0.42	0.60
	Item 5	0.54	0.52	0.62	0.44	0.57
	Item 6	0.62	0.62	0.60	0.58	0.66
Form D	Item 1	0.62	0.56	0.70	0.58	0.57
	Item 2	0.62	0.53	0.80	0.56	0.66
	Item 3	0.82	0.73	0.89	0.85	0.81
	Item 4	0.74	0.66	0.84	0.76	0.69
	Item 5	0.64	0.56	0.80	0.60	0.56
	Item 6	0.43	0.32	0.58	0.38	0.53
	Item 7	0.53	0.41	0.78	0.43	0.57

lecture-tutorial students exhibited small gains on Form A in previous semesters, cautions us from making any definitive statements at this point.

Finally, the differences in P-values on Forms B and C show a strikingly different pattern for the fall 2010 data than for either the fall 2009 or spring 2010 data. Whereas these differences suggested in previous semesters that the lecture-tutorial students were, in general, outperforming the non-lecture-tutorial students, the fall 2010 data show that the P-value changes for the lecture-tutorial students are equivalent to or even smaller than the changes for the non-lecture-tutorial students. This issue is re-addressed in upcoming sections.

Table 8.4: Cronbach's α for Forms A-D for fall 2010.

Form	Cronbach's α
Form A	0.50
Form B	0.59
Form C	0.50
Form D	0.77

8.2.2 Reliability

Table 8.4 gives Cronbach's α for each survey form for the fall 2010. Cronbach's α for Form D is significantly higher for this semester than for either of the previous semesters. This suggests that our expansion of and revisions to Form D should provide us a better handle on students' understandings of dark matter. Cronbach's α for Form B is approximately the same as it was for the spring 2010. The value of Cronbach's α decreased for both Form A and Form C. The reason for these drops is not immediately obvious. Form C is exactly the same as in the spring 2010 and we only revised a single question for Form A. These changes may be telling us our fall 2010 population is more homogeneous in its responses to these forms than in previous semesters. We have no hypothesis about why this might be the case.

8.2.3 Normalized Gains

Figure 8.1 shows the normalized gains for each item and for each form overall. Figure 8.1 echoes the results of our P-value analysis above. The normalized gains are much higher on Form D for the lecture-tutorial students than for the non-lecture-tutorial students. The lecture-tutorial students also show larger gains on items 1-5 on Form A than the non-lecture-tutorial students, although these gains never exceed 0.22. The normalized gains on Forms B and C are not, with a few exceptions, larger for the lecture-tutorial students than for the non-lecture-tutorial students. These gains suggest that the non-lecture-tutorial students did just as well as, if not slightly better than, the lecture-tutorial students on Forms B and C in the fall 2010. Of course, the relevance of all of these claims depends on whether or not the observed gains are statistically significant.

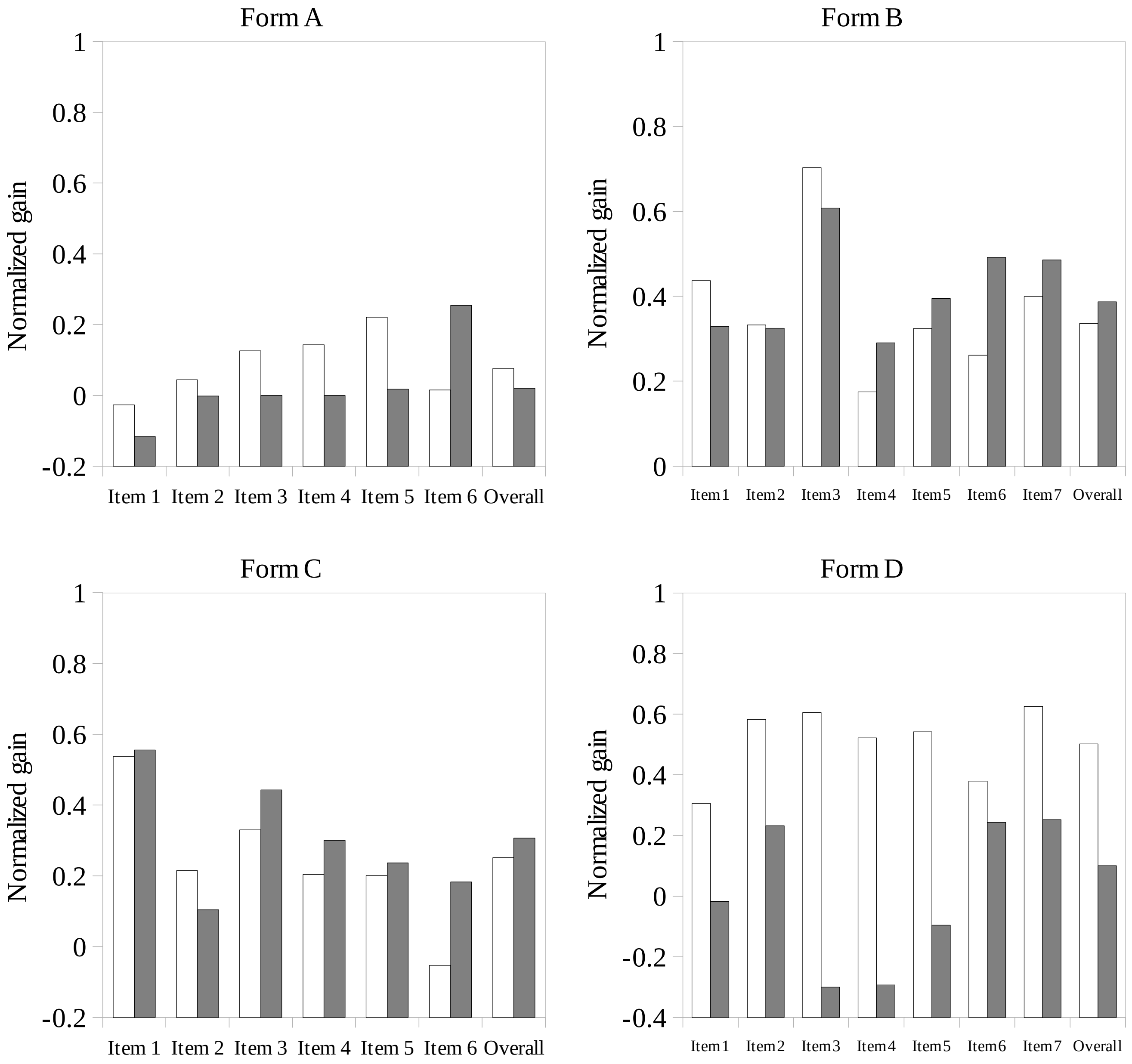


Figure 8.1: The normalized gains for Forms A-D for fall 2010. White bars correspond to the lecture-tutorial population and grey bars correspond to the non-lecture-tutorial population.

Table 8.5 shows the Mann-Whitney p-values for each form for both the lecture-tutorial and non-lecture-tutorial populations. In all but one case, the differences in the distributions of pre- and post-instruction scores for each population on each form is statistically significant ($p < 0.05$). The one exception is for the non-lecture-tutorial population on Form A. The fact that this group does

Table 8.5: Mann-Whitney p-values for the lecture-tutorial (LT) and non-lecture-tutorial (Non-LT) groups for Forms A-D for fall 2010.

Population	Form A	Form B	Form C	Form D
LT	0.0262	< 0.0001	< 0.0001	< 0.0001
Non-LT	0.1762	< 0.0001	< 0.0001	0.0150

not exhibit a statistically significant change in the distribution of scores is perhaps not surprising, given that Figure 8.1 shows they achieved normalized gains ≤ 0 on four of the six items. As in previous chapters, we will use item response theory (in Section 8.3.4 below) to check whether or not the differences in these gains are significant.

8.3 Item Response Theory Analysis

This section contains our IRT analysis of the response data from Forms A-C. We use Master's (1982) partial credit model, just as we did in previous chapters. Section 8.3.1 presents ConstructMap's estimates of the item's step difficulties and Thurstonian thresholds. Section 8.3.2 examines whether or not the data meet IRT's assumptions. Section 8.3.3 looks at how well the model fits the data. I discuss the reliability of Forms A-C from an IRT perspective in Section 8.3.4. Section 8.3.5 compares the IRT gains for the lecture-tutorial and non-lecture-tutorial populations.

8.3.1 Items' Difficulties

Table 8.6 gives the step difficulties and Thurstonian thresholds for the items on Forms A-C. In previous chapters we noted that some items were problematic when they had step difficulties and/or Thurstonian thresholds greater than 3 logits. Many of the values in Table 8.6 fall below or close to this limit. This suggests that our revisions have helped.

One notable exception to this pattern is item 2 on Form A. For this item, $b_{23} = 8.41$ logits and $\beta_3 = 8.41$ logits. These are extremely high, which implies that students have trouble achieving the highest score (3) on item 2. This result makes sense when one considers that item 2 asks about a contracting universe. Since all our lecture-tutorials focus on an expanding universe, this item

Table 8.6: The step difficulty b_{ij} and Thurstonian Threshold β_j parameters for the items on Forms A-C for fall 2010. All values are in logits.

		Step Parameters				Thurstonian Thresholds			
	Item	b_{01}	b_{12}	b_{23}	b_{34}	β_1	β_2	β_3	β_4
Form A	Item 1	-5.00	0.62	3.09	-	-5.01	0.55	3.17	-
	Item 2	-4.09	1.64	8.41	-	-4.09	1.64	8.41	-
	Item 3	-4.48	3.01	2.81	-	-4.48	2.47	3.35	-
	Item 4	-4.30	2.86	2.69	-	-4.30	2.33	3.22	-
	Item 5	-2.55	-0.92	-	-	-2.70	-0.77	-	-
	Item 6	-2.02	-0.19	0.15	-	-2.16	-0.45	0.56	-
Form B	Item 1	-1.46	-0.80	3.99	-1.38	-1.78	-0.52	1.33	1.39
	Item 2	-0.18	-1.70	2.61	-	-1.18	-0.73	2.62	-
	Item 3	-1.17	-0.36	-	-	-1.46	-0.07	-	-
	Item 4	-3.28	1.93	2.78	-	-3.29	1.66	3.06	-
	Item 5	-1.65	2.18	0.59	-	-1.67	1.19	1.62	-
	Item 6	-0.65	1.67	0.39	-	-0.76	0.89	1.34	-
	Item 7	-2.03	0.28	-	-	-2.12	0.37	-	-
Form C	Item 1	-3.02	1.36	0.34	-	-3.03	0.57	1.15	-
	Item 2	-1.79	1.40	2.14	-	-1.83	1.14	2.45	-
	Item 3	-0.98	2.49	-0.73	-	-1.03	0.86	1.04	-
	Item 4	-3.13	3.33	-0.90	-	-3.13	1.16	1.28	-
	Item 5	-2.57	2.56	-0.31	-	-2.58	1.02	1.26	-
	Item 6	-1.71	1.22	0.31	-	-1.77	0.52	1.09	-

requires students to apply their knowledge of Hubble plots to a new situation. The values of item 2's step difficulties and Thurstonian thresholds suggests that such an extension is difficult for many students.

8.3.2 Testing Item Response Theory's Assumptions

As in previous chapters, we look at whether or not our data meets the assumptions of item response theory. We again look at item outfit statistics (see Section 8.3.3 below) to inform us about the dimensionality of our surveys and Yen's Q3 statistic to tell us whether or not parameter invariance holds. We have cause for concern on both counts. In general, the outfit statistics reported below are smaller than expected, and three items fall outside the theoretically-expected bounds. Furthermore, Tables 8.7-8.9 show many item pairs for which Yen's Q3 statistic $> |0.20|$. Because we have evidence that the assumptions of IRT do not hold for our data, we will not claim that our data exhibits parameter invariance.

Table 8.7: Yen's Q3 statistic for each pair of items on Form A for the fall 2010.

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Item 1	1	**0.23**	0.06	0.05	-0.07	0.17
Item 2	**0.23**	1	0.14	0.16	-0.13	0.13
Item 3	0.06	0.14	1	**0.84**	0.03	**0.23**
Item 4	0.05	0.16	**0.84**	1	-0.02	**0.24**
Item 5	-0.07	-0.13	0.03	-0.02	1	**0.27**
Item 6	0.17	0.13	**0.23**	**0.24**	**0.27**	1

Table 8.8: Yen's Q3 statistic for each pair of items on Form B for the fall 2010.

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7
Item 1	1	0.10	-0.02	0.11	**0.26**	**0.27**	0.03
Item 2	0.10	1	0.05	-0.01	0.11	-0.06	0.05
Item 3	-0.02	0.05	1	**-0.23**	-0.13	-0.13	0.11
Item 4	0.11	-0.01	**-0.23**	1	0.15	0.13	-0.13
Item 5	**0.26**	0.11	-0.13	0.15	1	**0.24**	0.01
Item 6	**0.27**	-0.06	-0.13	0.13	**0.24**	1	0.17
Item 7	0.03	0.05	0.11	-0.13	0.01	0.17	1

Table 8.9: Yen's Q3 statistic for each pair of items on Form C for the fall 2010.

Item	Item 1	Item 2	Item 3	Item 4	Item 5	Item 6
Item 1	1	-0.11	0.13	0.08	0.14	0.14
Item 2	-0.11	1	0.11	0.09	0.07	-0.02
Item 3	0.13	0.11	1	**0.26**	**0.24**	0.06
Item 4	0.08	0.09	**0.26**	1	**0.32**	0.08
Item 5	0.14	0.07	**0.24**	**0.32**	1	0.19
Item 6	0.14	-0.02	0.06	0.08	0.19	1

8.3.3 Model Fit

Table 8.10 shows the outfit statistics for each item on Forms A-C. According to Wu and Adams (2011), we expect items on Form A to have outfit values between 0.85 and 1.15, items on Form B to have outfit values between 0.86 and 1.14, and items on Form C to have outfit values between 0.83 and 1.17. Three items have outfit values outside these bounds: item 5 on Form A

Table 8.10: Outfit statistics for the items on Forms A-C for the fall 2010.

Item	Form A	Form B	Form C
Item 1	1.03	0.97	1.04
Item 2	1.00	0.90	1.03
Item 3	0.99	**0.81**	0.90
Item 4	1.02	0.86	1.07
Item 5	**0.55**	0.90	0.99
Item 6	1.09	**0.78**	0.99
Item 7	-	0.86	-

and items 3 and 6 on Form B. We have no explanation for why these items have smaller outfit values than expected, especially since their outfit values for the spring 2010 were acceptable. In general, Table 8.10 shows that most items have outfit values less than unity, which implies that these items are actually *more* difficult than predicted by the partial credit model. This contrasts with the outfit values for the spring 2010 which implied that many of the items were *easier* than expected.

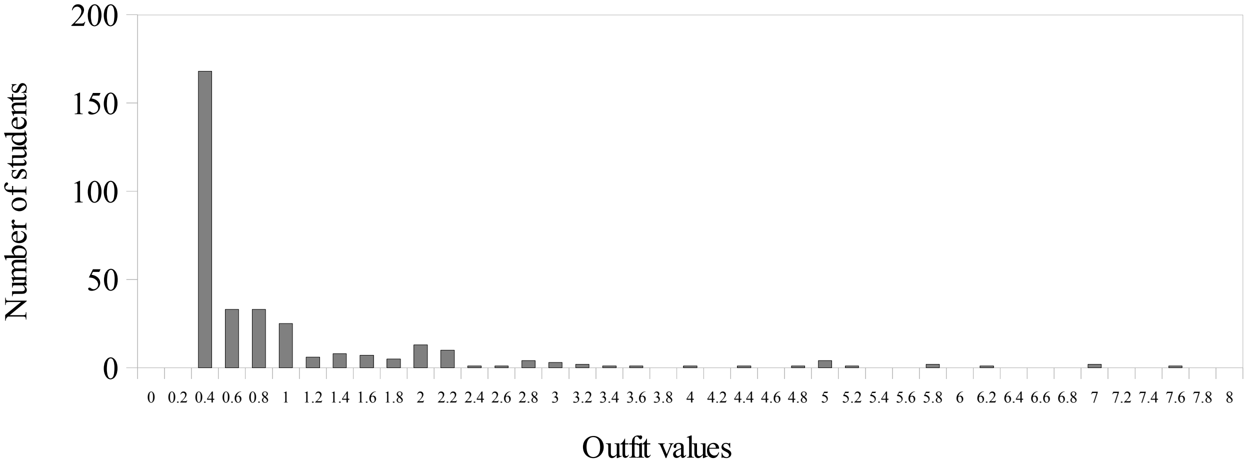


Figure 8.2: A histogram of students' outfit values for Form A for the fall 2010.

Figures 8.2-8.4 are histograms of respondents' outfit values. Wu and Adams (2011) predict that 95% of respondents to Forms A and C should have outfit values less than 2.13 and 95% of respondents to Form B should have outfit values less than 2.05. 91% of respondents to Form A have outfit values below 2.13. 90% of respondents to Form B have outfit values below 2.05. 96% of

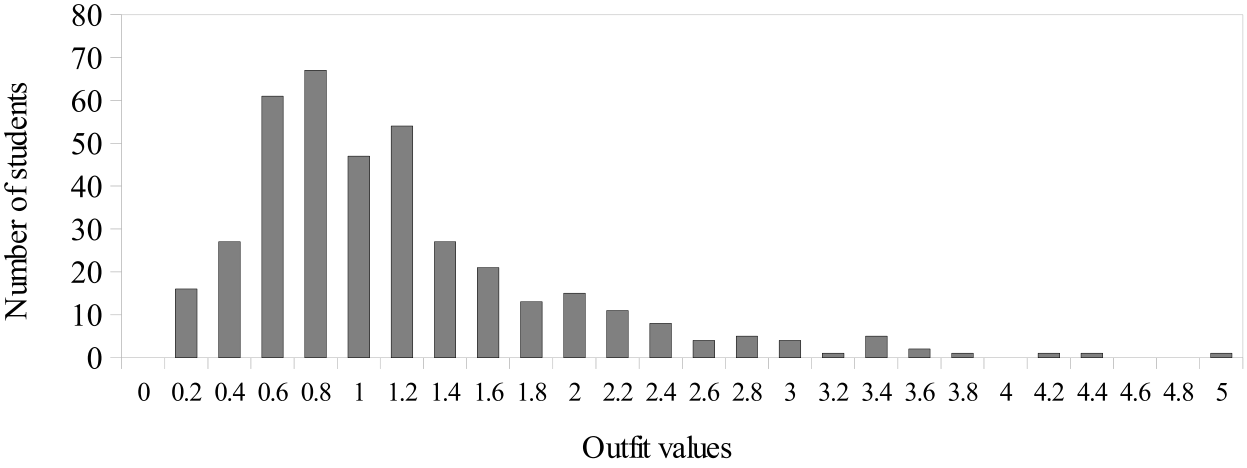


Figure 8.3: A histogram of students' outfit values for Form B for the fall 2010.

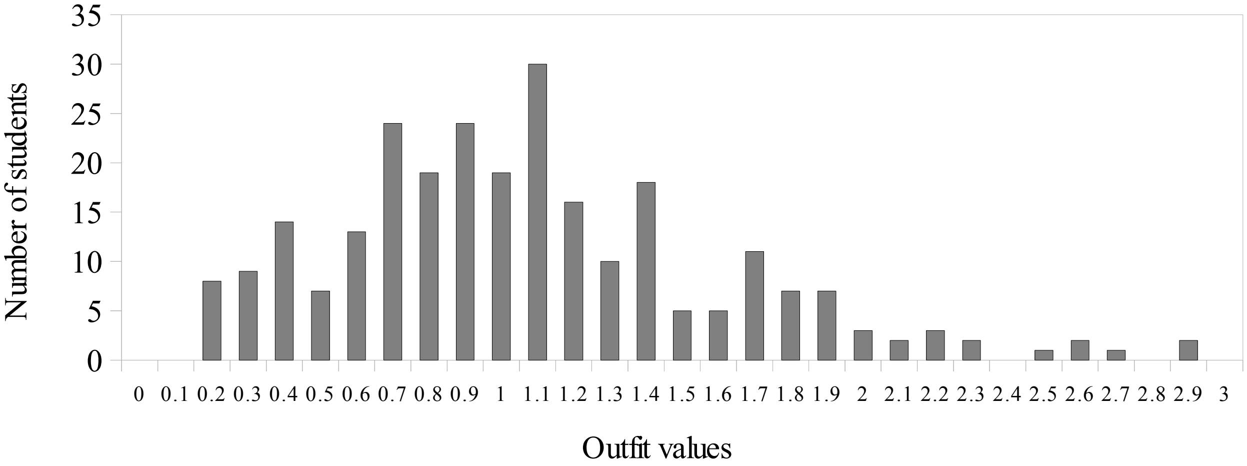


Figure 8.4: A histogram of students' outfit values for Form C for the fall 2010.

respondents to Form C have outfit values below 2.13. These values are close to those we reported for the spring 2010 data. They indicate that while the model fit is not horrible, it is also less than ideal.

8.3.4 Reliability

Are Forms A-C reliable? Figure 8.5 shows the standard error of measurement as a function of ability for Forms A-C and Figures 8.6-8.8 show the Wright Maps for Forms A-C, respectively.

The standard errors of measurement shown in Figure 8.5 are not as good as those for previous semesters. Nevertheless, the minima in the standard error of measurement plots roughly correspond to the peaks in students' abilities on the Wright maps shown in Figures 8.6-8.8. Furthermore, the Thurstonian thresholds of the items on each Form cover the range of observed abilities. This indicates that, from an IRT perspective, Forms A-C are reliable.

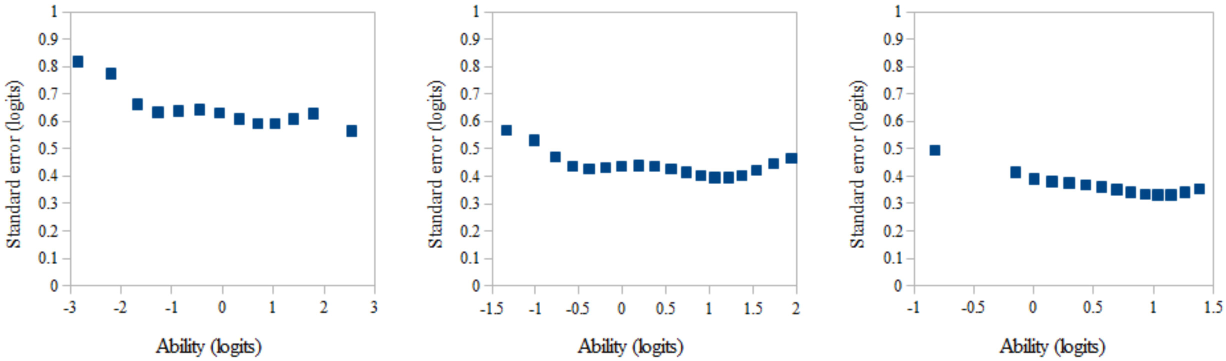


Figure 8.5: Standard error of measurement as a function of ability for (from left to right) Form A, Form B, and Form C for fall 2010.

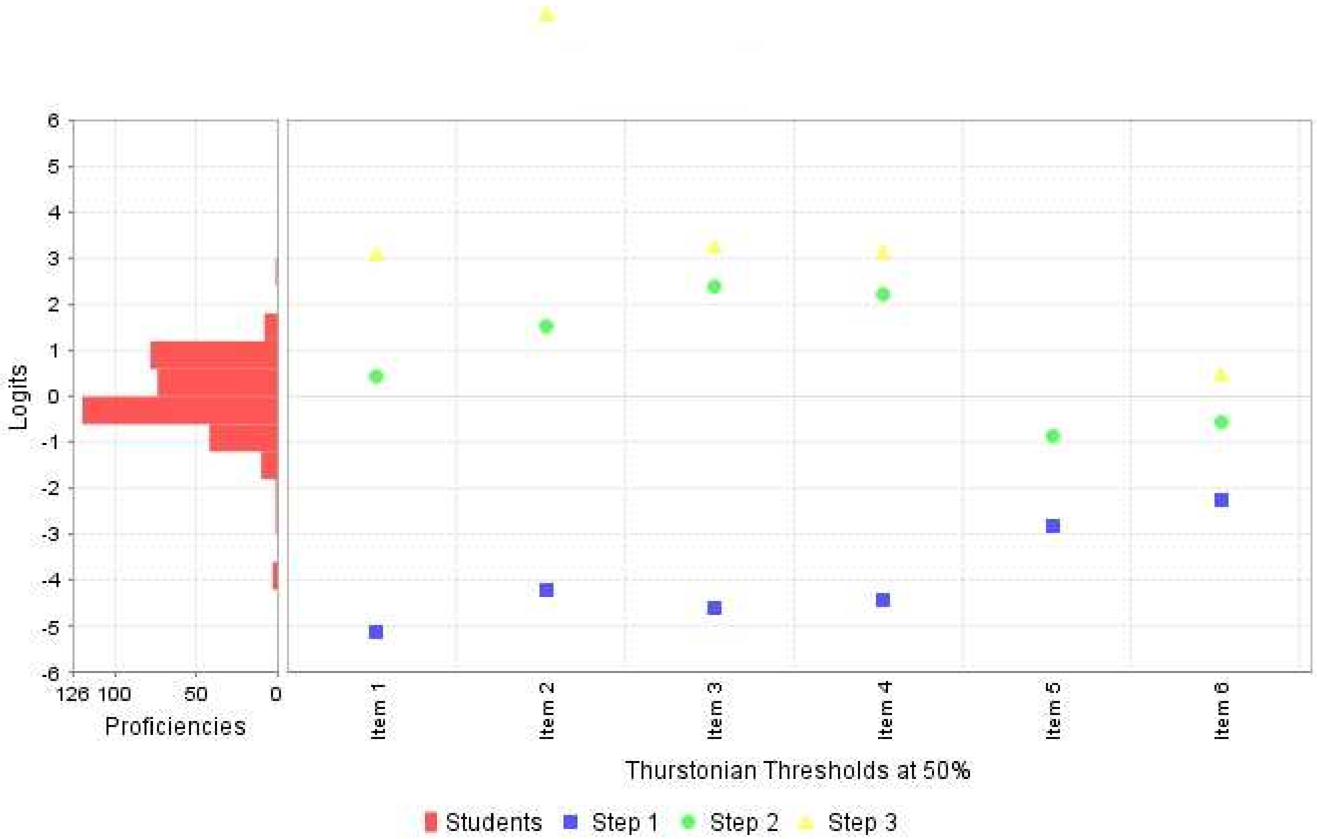


Figure 8.6: The Wright map for Form A for fall 2010.

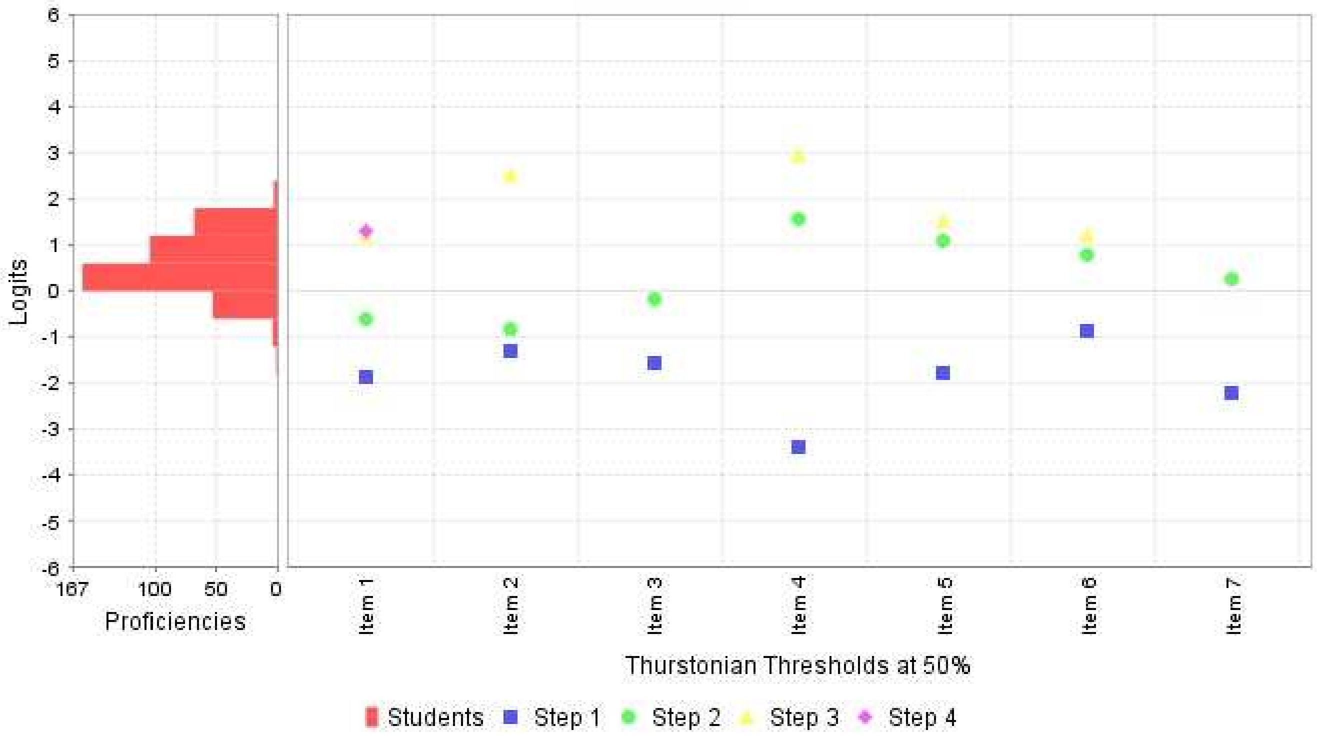


Figure 8.7: The Wright map for Form B for fall 2010.

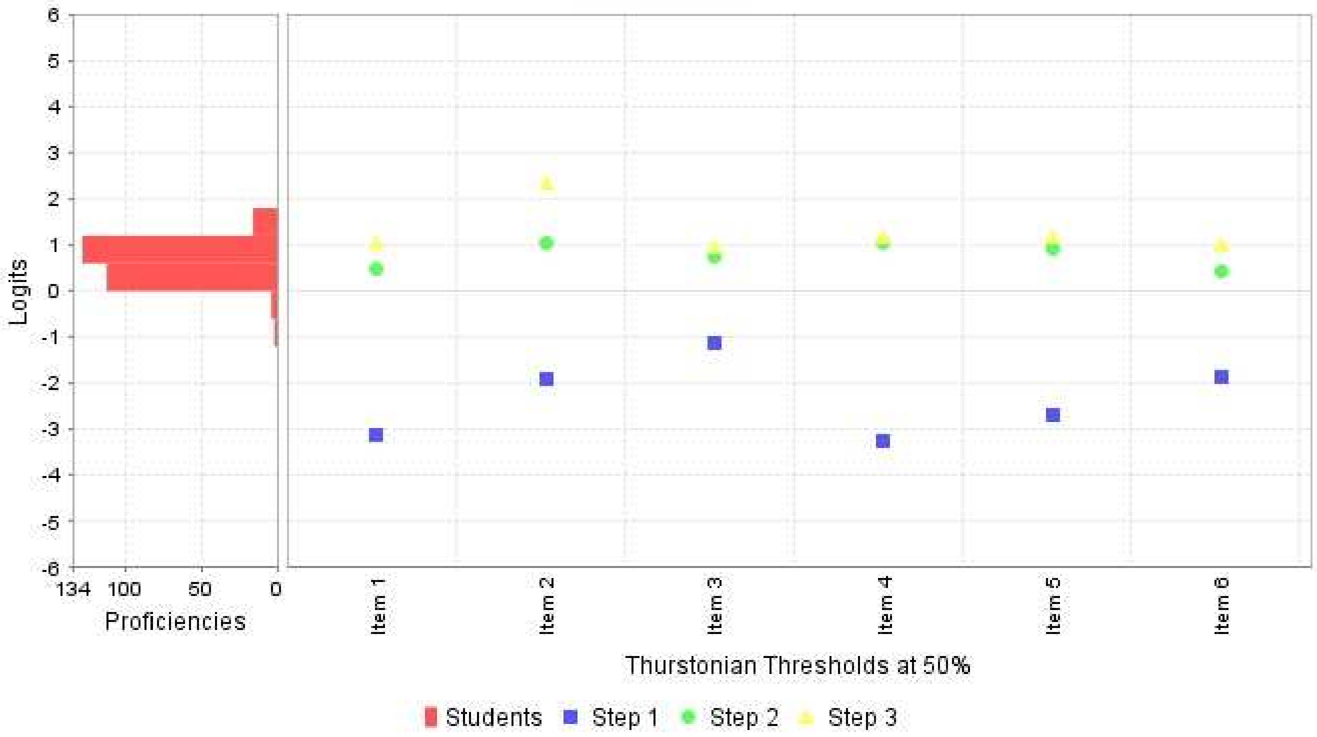


Figure 8.8: The Wright map for Form C for fall 2010.

Table 8.11: Average pre-instruction IRT scores, post-instruction IRT scores, and IRT gains for the lecture-tutorial and non-lecture-tutorial classes, as well as the difference between their gains, for Forms A-C in the fall 2010. All values are in logits.

Form	LT pre	Non-LT pre	LT post	Non-LT post	LT Gain	Non-LT Gain	LT Gain-Non-LT Gain
Form A	-0.07 ± 0.05	0.01 ± 0.09	0.27 ± 0.07	-0.06 ± 0.09	0.33 ± 0.08	-0.07 ± 0.12	0.41 ± 0.15
Form B	0.43 ± 0.03	0.34 ± 0.06	0.98 ± 0.04	1.01 ± 0.06	0.55 ± 0.05	0.67 ± 0.08	-0.12 ± 0.10
Form C	0.59 ± 0.04	0.48 ± 0.05	0.85 ± 0.04	0.85 ± 0.05	0.27 ± 0.06	0.37 ± 0.07	-0.10 ± 0.09

8.3.5 IRT Gains

How do the IRT-estimated abilities of the lecture-tutorial students compare to the estimated abilities of the non-lecture-tutorial students? Table 8.11 contains the answer. According to Table 8.11, the IRT gains (post- minus pre-instruction average ability) are larger for the non-lecture-tutorial students on Form B and C. We cannot claim that the lecture-tutorial students showed a greater improvement on Forms B and C than their non-lecture-tutorial counterparts.

Form A is a different story. In previous semesters, we saw no difference in the gains of lecture-tutorial and non-lecture-tutorial students. For the fall 2010, the gain for the lecture-tutorial students is 0.41 ± 0.15 logits larger than for the non-lecture-tutorial students. This suggests that the revisions we made to the "Hubble's Law" lecture-tutorial between the spring and fall 2010 semesters may have contributed to the lecture-tutorial students' improved average performance.

8.4 Breakdown of Item Responses

Our CTT and IRT analyses of students' responses indicate that the lecture-tutorial students achieved larger learning gains than the non-lecture-tutorial students on Forms A and D in the fall of 2010. The same is not true for Forms B and C. On these two forms, we have no evidence of larger learning gains for the lecture-tutorial students; if anything, the data suggest that the non-lecture-tutorial students did better. These results warrant further inspection.

We need to examine students' responses and reasoning patterns for several reasons. First, previous semesters have shown little difference between the responses of lecture-tutorial students to Form A and the responses of non-lecture-tutorial students. The fact that lecture-tutorial students in the fall 2010 have higher learning gains than their non-lecture-tutorial peers suggests there is something different about their responses for this semester. Second, data from previous semesters shows the lecture-tutorial students outperforming the non-lecture-tutorial students on Forms B and C. The fact that the fall 2010 data do not also exhibit this pattern demands explanation. Third, we substantially revised Form D prior to the fall of 2010 in order to give us a better idea of common student difficulties related to dark matter. We thus need to look at students' responses to see if we can detect such difficulties. In general, the data presented in this section further illuminates our understanding of students' difficulties with cosmological topics.

8.4.1 Form A Responses

Table 8.12 shows the distribution of scores for both the lecture-tutorial and non-lecture-tutorial students, pre- and post-instruction, on Form A. Unlike previous semesters, Table 8.12 shows a greater percent of lecture-tutorial students earned scores of 2 and 3 post-instruction than they did pre-instruction. In contrast, the percent of non-lecture-tutorial students earning scores of 2 and 3 stayed constant pre- to post-instruction: For both items, 2% received a 2 and 0% received a 3. Furthermore, while 0% of all respondents received a 3 on item 2 (pre- and post-instruction), the percent of lecture-tutorial students with a score of 2 increased pre- to post-instruction, while the percent with scores of 0 and 1 decreased. The distribution of scores for the non-lecture-tutorial students on item 2 did not change much pre- to post-instruction. These trends help explain why the lecture-tutorial students exhibited larger learning gains than the non-lecture-tutorial students on Form A in the fall 2010.

We can see why the lecture-tutorial students did better than the non-lecture-tutorial students by looking at their graph selections for items 1-4. Figures 8.9-8.12 show the percent of students who chose each graph pre- and post-instruction for both populations of students. Neither the lecture-tutorial nor the non-lecture-tutorial students show much of a change in their graph selections for item 1: Graphs C and F are the most popular choices pre- and post-instruction. However, the situation is different on items 2-4. There is a larger, positive change in the percent of lecture-

Table 8.12: Distribution of scores on Form A for fall 2010.

	LT Pre				Non-LT Pre			
	0	1	2	3	0	1	2	3
Item 1	1%	61%	35%	2%	0%	51%	37%	12%
Item 2	4%	79%	17%	0%	0%	78%	22%	0%
Item 3	2%	94%	4%	0%	2%	96%	2%	0%
Item 4	2%	93%	5%	0%	2%	96%	2%	0%
Item 5	6%	27%	67%	-	4%	33%	63%	-
Item 6	10%	26%	25%	39%	10%	39%	24%	27%
	LT Post				Non-LT Post			
	0	1	2	3	0	1	2	3
Item 1	0%	67%	32%	1%	2%	53%	43%	2%
Item 2	1%	76%	23%	0%	2%	74%	23%	0%
Item 3	1%	76%	18%	5%	2%	96%	2%	0%
Item 4	1%	73%	19%	6%	2%	96%	2%	0%
Item 5	3%	26%	72%	-	6%	28%	66%	-
Item 6	6%	26%	35%	33%	4%	26%	34%	36%

tutorial students choosing the right graph for items 2-4 (graph B for item 2, graph A for item 3, and graph D for item 4) than there is for the non-lecture-tutorial students. That's not to say that an overwhelming percent of lecture-tutorial students chose the right graph post-instruction on these items. On items 3 and 4, for example, the correct graphs are always chosen by a minority of students. Figures 8.11 and 8.12 show that this becomes a larger minority for the lecture-tutorial students post-instruction.

What are we to make of these results? We did not uncover any new, widespread conceptual difficulty or reasoning errors in students' interpretations of Hubble plots. The students in our fall 2010 sample are still experiencing the same difficulties as students in the fall 2009 and spring 2010 samples. However, the above data suggest that a greater percent of lecture-tutorial students post-instruction are selecting the right answers and are articulating the right reasons for their selections than are their non-lecture-tutorial counterparts. This suggests that our most recent revisions to the cosmology lecture-tutorials, especially the "Hubble's Law" lecture-tutorial, are helping students improve their abilities to interpret Hubble plots.

What about items 5 and 6 on Form A? Table 8.12 shows improvements in the distribution of

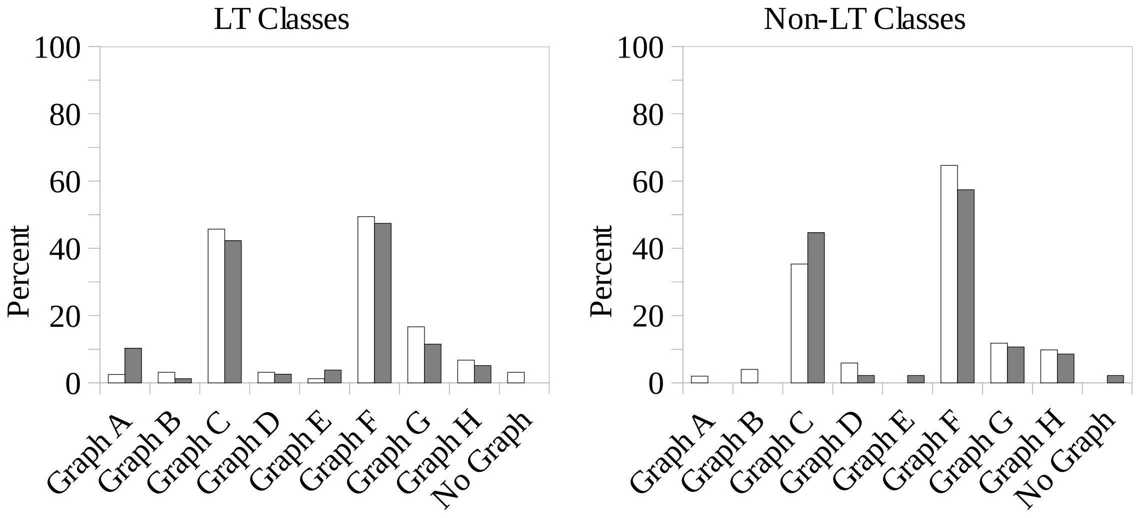


Figure 8.9: Students' graph choices for item 1 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

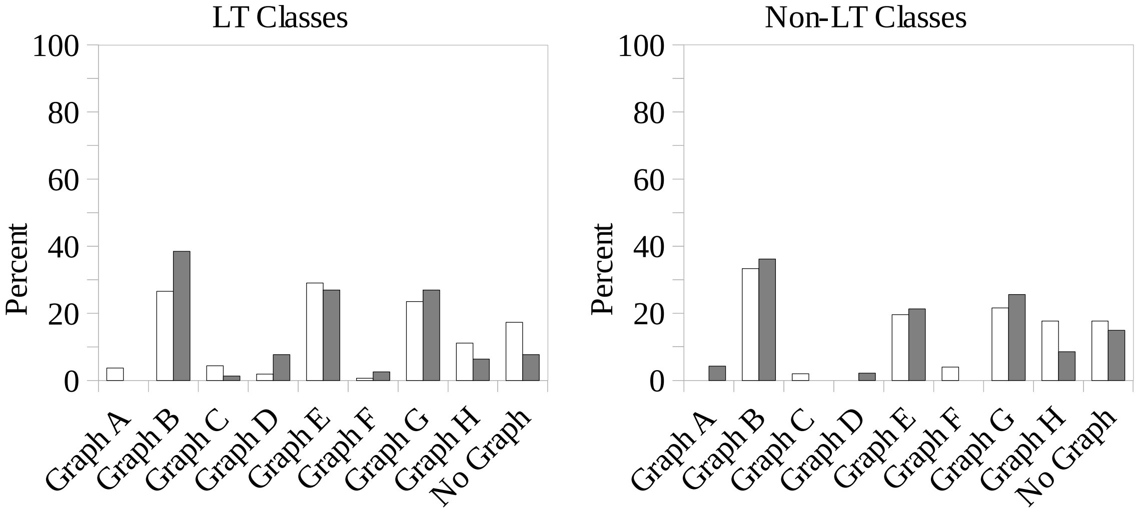
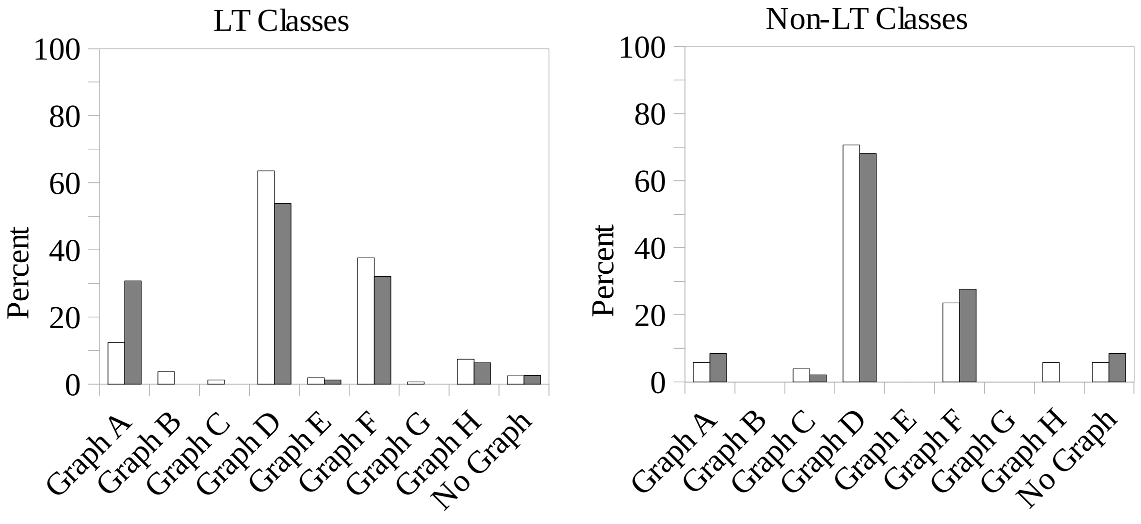


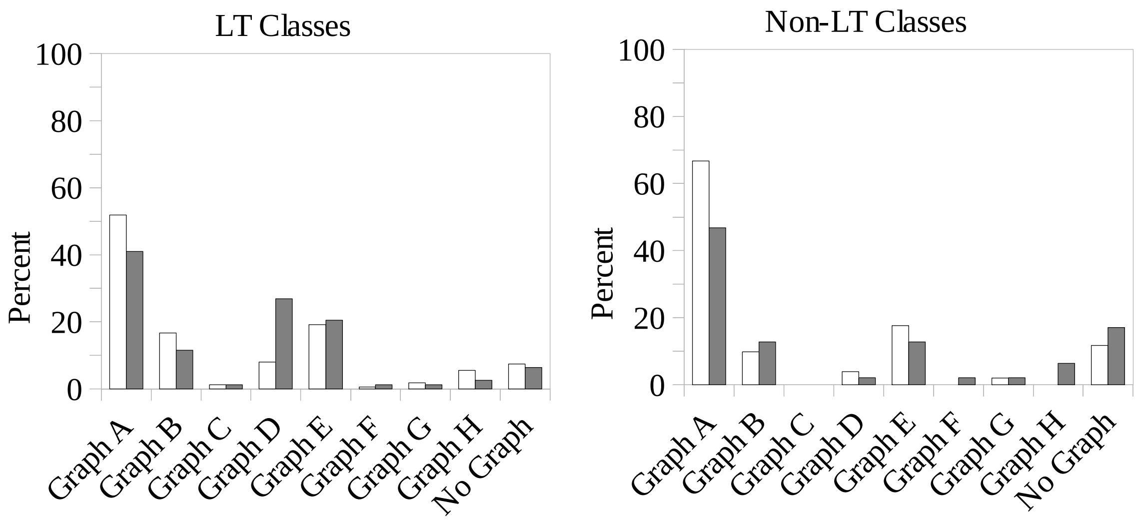
Figure 8.10: Students' graph choices for item 2 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

scores for item 5 for both populations of students, although the change in the percent of students earning a 2 is greater for the lecture-tutorial group. The Hake plot in Figure 8.1 also shows a larger normalized gain for the lecture-tutorial students than for the non-lecture-tutorial students on this item. Table 8.13 shows the reasoning elements used by students in their responses to item 5.


LT Classes

Non-LT Classes

Figure 8.11: Students' graph choices for item 3 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.


LT Classes

Non-LT Classes

Figure 8.12: Students' graph choices for item 4 on Form A. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

A greater percent of lecture-tutorial students (pre- and post-instruction) than non-lecture-tutorial students claim that a steeper slope corresponds to a faster expansion rate. Otherwise, there are not many differences in the percent of students using a particular reasoning element when one compares the pre-instruction responses of both groups. The same is true when one compares the post-instruction responses of both groups.

Table 8.12 and Figure 8.1 suggest that the non-lecture-tutorial students have larger learning gains than the lecture-tutorial students on item 6. Table 8.14, which shows the reasoning elements used by students answering this item, suggests a couple of reasons why. First, the change in the percent of students drawing or discussing a flatter slope and talking about a decreased velocity or speed is greater for the non-lecture-tutorial students. In fact, the percent changes for the lecture-tutorial students are negative for both of these elements. Second, the percent of lecture-tutorial students who (incorrectly) draw or discuss a non-constant slope in their answer increases pre- to post-instruction, whereas this percent decreases for the non-lecture-tutorial students. These patterns explain why the non-lecture-tutorial students did better than the lecture-tutorial students on item 6.

8.4.2 Form B Responses

Table 8.15 shows the distribution of scores on Form B for the fall 2010. A cursory examination of Table 8.15 shows that, for several items, the changes in the distribution of scores pre- to post-instruction for the lecture-tutorial students mirror the changes in the distribution of scores for the non-lecture-tutorial students. This seems to contradict the normalized gains in Figure 8.1 which show the non-lecture-tutorial students posting higher gains than the lecture-tutorial students on a majority of Form B's items. This further underscores the need to look at the pattern of students' responses to each item in detail.

Item 1 is an example of an item for which the lecture-tutorial students achieved a higher normalized gain than the non-lecture-tutorial students. This result makes sense in light of the distribution of scores shown in Table 8.15: While both populations have a similar pre-instruction distribution of scores, 44% of the lecture-tutorial students post-instruction earned the maximum score of 4 compared to 29% of the non-lecture-tutorial students. This is further explained when one looks at Table 8.16, which shows the reasoning elements used by both populations of students pre- and post-instruction. The largest difference between the two populations is in the percent who talked about the movement/increasing distances of galaxies. For the lecture-tutorial students,

Table 8.13: Common reasoning elements used by students in their answers to Form A, Item 5: Use the blank graph provided below to draw what your think Figure 1 would look like if the universe had been expanding twice as fast. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student draws or discusses a steeper slope.	70%	73%	65%	68%
Student draws or discusses a flatter slope.	9%	8%	8%	2%
Student draws or discusses an unchanged slope.	1%	3%	10%	4%
Student draws a line with a non-constant slope.	8%	13%	8%	11%
Student draws a line whose slope is ≤ 0.	2%	0%	2%	0%
Student draws a line whose y-intercept $\neq 0$.	7%	4%	6%	4%
Student talks about increased velocity.	39%	35%	35%	47%
Student talks about decreased velocity.	1%	1%	0%	0%
Student talks about velocity staying the same.	1%	1%	2%	2%
Student talks about increased distance.	11%	13%	8%	17%
Student talks about decreased distance.	4%	4%	0%	0%
Student talks about distance staying the same.	4%	1%	6%	6%
Student compares variable x to variable y by saying x is more than y.	2%	0%	0%	0%
Student says distance and velocity change by the same amount.	2%	0%	2%	4%
Student says s/he can't answer the question without a time axis/variable.	1%	0%	2%	0%
Student gives irrelevant information.	0%	0%	0%	0%
Student gives some other reason not specified above.	5%	5%	2%	9%
Student has no idea.	3%	0%	8%	4%
Answer field is blank.	4%	3%	2%	0%

Table 8.14: Common reasoning elements used by students in their answers to Form A, Item 6: 6) Use the blank graph provided below to draw what you think Figure 1 would look like for our universe if it took much longer to reach its current size. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student draws or discusses a positive slope.	83%	88%	80%	85%
Student draws or discusses a negative slope.	4%	3%	4%	0%
Student draws or discusses a slope of zero.	1%	0%	2%	0%
Student draws or discusses a non-constant slope.	12%	22%	18%	15%
Student draws or discusses a steeper slope.	5%	5%	4%	0%
Student draws or discusses a flatter slope.	66%	64%	57%	70%
Student draws or discusses an unchanged slope.	1%	1%	4%	2%
Student draws a line whose y-intercept $\neq 0$.	6%	1%	8%	2%
Student talks about a faster expansion rate.	0%	0%	0%	0%
Student talks about a slower expansion rate.	12%	14%	10%	19%
Student talks about an unchanged expansion rate (i.e. expansion rate is the same as Figure 1).	0%	0%	0%	0%
Student talks about an expansion rate that speeds up.	1%	0%	0%	0%
Student talks about an expansion rate that slows down.	1%	0%	0%	0%
Student talks about a constant expansion rate (i.e. expansion rate doesn't change over time).	0%	0%	0%	0%
Student talks about an increased velocity/speed.	2%	4%	4%	2%
Student talks about a decreased velocity/speed.	39%	36%	29%	36%
Student talks about a constant velocity/speed.	1%	1%	2%	2%
Student talks about an increased distance.	7%	5%	8%	4%
Student talks about a decreased distance.	2%	6%	4%	9%
Student talks about a constant distance.	2%	1%	0%	0%
Student talks about what happened to the universe in the past.	0%	0%	0%	0%
Student talks about what will happen to the universe in the future.	0%	0%	0%	0%
Student says there's not enough information.	3%	1%	10%	9%
Students says the universe's age is irrelevant.	0%	0%	0%	0%
Student talks about the time the universe needs to reach its current size.	0%	4%	0%	4%
Student says s/he needs to know how the expansion rate changes.	0%	0%	0%	0%
Student says there's no time variable/axis.	1%	0%	2%	4%
Student says the significance of the universe's age is unclear.	0%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%
Student gives some other reason not specified above.	5%	5%	4%	4%
Student has no idea.	4%	0%	12%	4%
Answer field is blank.	7%	6%	2%	0%

Table 8.15: Distribution of scores on Form B for fall 2010.

	LT Pre					Non-LT Pre				
	0	1	2	3	4	0	1	2	3	4
Item 1	5%	25%	55%	1%	13%	4%	28%	54%	2%	13%
Item 2	8%	13%	74%	5%	-	11%	9%	76%	4%	-
Item 3	11%	33%	56%	-	-	9%	31%	59%	-	-
Item 4	2%	84%	13%	1%	-	2%	87%	11%	0%	-
Item 5	7%	74%	7%	12%	-	11%	78%	6%	6%	-
Item 6	15%	53%	16%	16%	-	28%	46%	13%	13%	-
Item 7	3%	50%	47%	-	-	7%	50%	43%	-	-
	LT Post					Non-LT Post				
	0	1	2	3	4	0	1	2	3	4
Item 1	2%	2%	51%	1%	44%	2%	2%	59%	8%	29%
Item 2	5%	1%	65%	29%	-	0%	4%	78%	18%	-
Item 3	0%	16%	84%	-	-	2%	16%	82%	-	-
Item 4	3%	55%	35%	7%	-	0%	49%	37%	14%	-
Item 5	10%	34%	21%	35%	-	8%	35%	24%	33%	-
Item 6	18%	29%	11%	42%	-	10%	24%	20%	47%	-
Item 7	5%	23%	71%	-	-	4%	25%	71%	-	-

this difference went from 26% pre-instruction to 49% post-instruction. For the non-lecture-tutorial students, this went from 17% to 31%. The fact that this pre- to post-instruction change was greater for the lecture-tutorial students than for the non-lecture-tutorial students explains the larger normalized gain on item 1 for the lecture-tutorial students.

Table 8.16 also shows many of the same difficulties uncovered in previous semesters. For example, 13% of the lecture-tutorial population and 11% of the non-lecture-tutorial population pre-instruction claimed that “the expansion of the universe” is a metaphor for humans learning more about the universe over time. Another 12% of the lecture-tutorial students and 20% of the non-lecture-tutorial students said expansion refers to the creation of new objects in the universe over time. The percent of students in these populations making either of these claims dropped to 0%-2% post-instruction.

According to Figure 8.1, the normalized gains for lecture-tutorial and non-lecture-tutorial students on item 2 are approximately the same. In previous semesters, the lecture-tutorial students achieved much higher gains on this item than the non-lecture-tutorial students. One might suspect

Table 8.16: Common reasoning elements used by students in their answers to Form B, Item 1: Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding."

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says the size of the universe increases over time.	28%	29%	24%	27%
Student says the universe has a center.	7%	3%	7%	0%
Student says the universe has no center.	0%	1%	0%	4%
Student says the universe has an edge.	3%	0%	4%	0%
Student says the universe has no edge.	0%	2%	2%	4%
Student talks about redshifts/Doppler shifts.	2%	1%	0%	6%
Student says space(time) is growing/stretching.	5%	38%	4%	45%
Student talks about the movement of galaxies and/or their increasing distances.	26%	49%	17%	31%
Student talks about the movement of stars and/or their increasing distances.	14%	3%	7%	4%
Student talks about the movement of planets and/or their increasing distances.	5%	2%	6%	4%
Student talks about the movement of objects (something unspecified or not a star, planet, or galaxy) and/or their increasing distances.	16%	12%	15%	4%
Student says the distances between everything increase.	8%	10%	9%	14%
Student says farther objects move away faster.	2%	8%	4%	12%
Student talks about the Big Bang.	12%	8%	24%	10%
Student talks about an explosion.	2%	1%	7%	0%
Student says the early universe was once hot, small, and/or dense.	4%	8%	4%	4%
Student says we learn more about the universe over time.	13%	2%	11%	0%
Student talks about how we are looking further back in time as we look farther into space.	4%	0%	0%	0%
Student says new things are created in the universe over time.	12%	1%	20%	2%
Student gives irrelevant information.	0%	0%	0%	0%
Student gives some other reason not specified above.	1%	2%	13%	4%
Student has no idea.	1%	0%	4%	2%
Answer field is blank or the student provided no reason or explanation.	4%	2%	4%	2%

that this should also be the case for the fall 2010 semester after examining Table 8.15, which shows a greater shift toward higher scores for the lecture-tutorial students than for the non-lecture-tutorial students. This is further supported by the information in Table 8.17, which shows the reasoning elements used by students in their answers to item 2. The percent of lecture-tutorial students claiming the Big Bang marked the beginning of expansion went from 23% pre-instruction to 51% post-instruction, while for the non-lecture-tutorial students it went from 13% pre-instruction to 27% post-instruction. Likewise, the percent of lecture tutorial students saying the Big Bang was an explosion dropped from 55% to 23%, while for non-lecture-tutorial students the drop was from 57% to 35%. Finally, the percent of students who claim that matter existed prior to the Big Bang decreased for both populations, although the non-lecture-tutorial population showed the larger decrease. Overall, the results in Table 8.17 imply that the lecture-tutorials are helping students.

So why do the non-lecture-tutorial students have almost the same normalized gain on this item as the lecture-tutorial students? There are three related reasons. First, the percent of non-lecture-tutorial students with a score of 0 on item 2 went from 11% pre-instruction to 0% post-instruction. This shift no doubt boosted the normalized gain for the non-lecture-tutorial population on this item. Second, a smaller percent of non-lecture-tutorial students wrote (post-instruction) that the Big Bang was an explosion than in previous semesters. Previously, the percent of non-lecture-tutorial students making such a claim remained above 50% post-instruction. Finally (and as mentioned above), the non-lecture-tutorial students exhibit a greater decrease in the percent of students saying that matter existed before the Big Bang than the lecture-tutorial students. These factors combined helped the normalized gains of the lecture-tutorial and non-lecture-tutorial students appear approximately equal on item 2, despite the fact that Tables 8.15 and 8.17 indicate a greater improvement in the overall performance for the lecture-tutorial students.

Figure 8.1 shows that the lecture-tutorial students have a larger normalized gain on item 3 than the non-lecture-tutorial students. This results is somewhat puzzling when one looks at the distribution of scores in Table 8.15. The distribution is roughly the same for both populations pre-instruction and it is also roughly the same for both populations post-instruction. Overall, the

Table 8.17: Common reasoning elements used by students in their answers to Form B, Item 2: Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory."

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says the Big Bang is the beginning of the universe/everything in universe.	43%	46%	50%	73%
Student says the Big Bang is the beginning of expansion.	23%	51%	13%	27%
Student says the Big Bang was the beginning of something smaller than the universe.	9%	1%	11%	4%
Student says the Big Bang is an event that happened to something smaller than the universe.	3%	0%	0%	2%
Student says the Big Bang is the beginning of space.	1%	3%	2%	16%
Student says the Big Bang is the beginning of time.	1%	3%	0%	16%
Student talks about the creation/production of elements.	3%	6%	4%	4%
Student says the Big Bang was an explosion.	55%	23%	57%	35%
Student says the Big Bang was not an explosion.	0%	1%	0%	0%
Student says matter existed before the Big Bang.	33%	26%	48%	25%
Student says there was a dense piece of matter before the Big Bang.	13%	8%	13%	12%
Student talks about matter coming together before the Big Bang.	9%	4%	2%	4%
Student says matter formed from energy.	0%	15%	2%	8%
Student says the early universe was hot, dense, and/or small.	9%	38%	6%	22%
Student says the Big Bang was an event that happened in empty space.	1%	4%	4%	0%
Student gives irrelevant information.	1%	1%	0%	2%
Student gives some other reason not specified above.	3%	2%	2%	0%
Student says s/he has no idea.	3%	0%	7%	0%
Answer field is blank or the student provided no reason.	6%	4%	4%	0%

increase in the percent of students with a score of 2 pre- to post-instruction is slightly greater for the lecture-tutorial students than for the non-lecture-tutorial students. Similarly, the decrease in the percent of students with scores of 0 and 1 pre- to post-instruction is slightly greater for the lecture-tutorial students than for the non-lecture-tutorial students. Table 8.18 shows the reasoning elements students used. There are three interesting patterns to note in this table. First, the change in the percent of students saying they should see more galaxies is greater for the lecture-tutorial population than for the non-lecture-tutorial population. This is also true for the percent of students saying that Galaxy X's observable universe extends into regions beyond our observable universe. Finally, there was a greater decrease pre- to post-instruction in the percent of lecture-tutorial students compared to non-lecture-tutorial students saying inhabitants of Galaxy X will see nothing/blackness/empty space. The combined effect of all of these facts likely explains the larger normalized gain for the lecture-tutorial students on this item.

Item 4 is an example of an item for which the non-lecture-tutorial students achieved a higher learning gain than the lecture-tutorial students. It is also an item that we revised prior to the fall 2010 semester. Figure 8.13 shows the percent of students who selected choices A, B, and C in item 4. While a greater percent of both populations post-instruction select the correct answer (C) than pre-instruction, this change is larger for the non-lecture-tutorial population. Table 8.19 shows the reasoning elements invoked by students for each of the three answer choices. Both populations show an increase in the percent of students who say there was no space before the Big Bang/no space outside of the Big Bang/all of space was part of the Big Bang, an increase in the percent of students who say matter did not exist before the Big Bang, and a decrease in the percent of students who say matter existed before the Big Bang. However, these changes are larger for the non-lecture-tutorial population. Thus, not only did a greater percent of the non-lecture-tutorial students chose the correct answer post-instruction, but they supported their choices with correct reasons. This explains why the non-lecture-tutorial population has a larger learning gain on this item than the lecture-tutorial population.

Aside from the issue of which group achieved the larger learning gain, Figure 8.13 also raises

Table 8.18: Common reasoning elements used by students in their answers to Form B, Item 3: Each dot in the picture on the left is a galaxy. The Milky Way Galaxy (the one we live in) is at the center of the picture. All of the galaxies inside the circle can be seen from Earth. Any galaxies that exist outside the circle are so far away that their light has not had time to reach Earth. Describe what inhabitants of Galaxy X probably see when they look in the direction of the arrow.

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says they'd see more galaxies.	45%	58%	56%	63%
Student says they'd see more stars.	13%	11%	9%	2%
Student says they'd see more planets.	3%	1%	0%	0%
Student says they'd see more objects (otherwise unspecified).	2%	1%	0%	0%
Student says they'd see something similar to what we see.	19%	47%	24%	51%
Student says they'd see things we cannot see.	22%	29%	22%	22%
Student says they'd see nothing/blackness/empty space.	23%	7%	17%	8%
Student says they'd see objects separated by greater distances than what we see. .	1%	0%	0%	0%
Students say Galaxy X's observable universe extends into regions outside of our observable universe.	21%	40%	17%	29%
Student gives irrelevant information.	4%	1%	4%	0%
Student gives some other reason not specified above.	4%	2%	9%	2%
Student has no idea.	1%	0%	4%	0%
Answer field is blank.	7%	0%	2%	2%

another point. Choice A, which essentially claims that the Big Bang was an explosion of pre-existing matter into empty space, is the most popular choice pre-instruction and is still selected by a significant minority post-instruction for both populations. This lends further support to previous papers (Lineweaver and Davis 2005; Prather, Slater, and Offerdahl 2003) that claim this is a common difficulty experienced by many people.

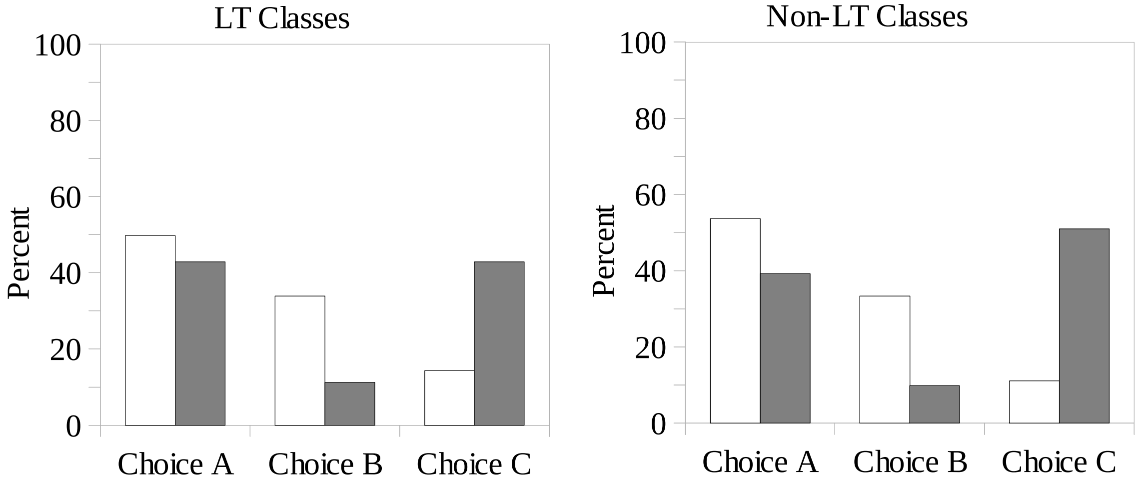


Figure 8.13: Percent of students who selected choice A, B, or C in response to item 4 on Form B. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Item 5 is another item for which the non-lecture-tutorial students have a larger learning gain than the lecture-tutorial students. Table 8.20 shows the reasoning elements used by students in response to this item. Pre- to post-instruction, a greater percent of students in both populations say the universe is infinite/has no edges and the universe is the same everywhere. Pre- to post-instruction, there is a decrease in the percent of students who say that the center of the universe is where the Big Bang happened/where the universe began/where the universe is expanding from. A significant difference in the lecture-tutorial and non-lecture-tutorial populations is in the percent who claim that the universe has no center or the center changes because of expansion/objects are in motion. For the lecture-tutorial population, this percent increases pre- to post-instruction, while for the non-lecture-tutorial population it decreases. This helps explain why the non-lecture-tutorial population has the larger normalized gain on this item.

Table 8.19: Common reasoning elements used by lecture-tutorial students in their answers to Form B, Item 4: Circle the sentence that best describes the universe at the time of the Big Bang: a) In the beginning, there was space in the universe surrounding the location of the Big Bang but this space was empty of all matter. b) In the beginning, there was space in the universe surrounding the location of the Big Bang and matter already existed in this space. c) I think of the Big Bang differently than a or b. Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	LT Pre			LT Post			Non-LT Pre			Non-LT Post		
Reasoning Element	A	B	C	A	B	C	A	B	C	A	B	C
Student says space existed before the Big Bang.	11%	5%	4%	7%	0%	5%	7%	6%	0%	15%	0%	4%
Student says there was no space before the Big Bang/no space outside of the Big Bang/all of space was part of the Big Bang.	1%	0%	33%	0%	0%	59%	0%	0%	17%	0%	0%	73%
Student says space outside the Big Bang is necessary for the Big Bang.	0%	2%	0%	2%	0%	0%	0%	0%	17%	0%	0%	0%
Student says matter existed before the Big Bang.	15%	31%	37%	12%	18%	24%	3%	22%	50%	10%	20%	15%
Student says matter did not exist before the Big Bang.	11%	0%	11%	24%	0%	22%	14%	0%	0%	30%	0%	27%
Student says matter existing before the Big Bang is necessary for the Big Bang.	3%	14%	4%	2%	9%	7%	0%	17%	0%	0%	0%	0%
Student says matter cannot be created or destroyed.	0%	5%	4%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says all matter was once in one spot.	4%	3%	4%	7%	9%	15%	3%	6%	17%	5%	0%	8%
Student says expansion = matter filling empty space.	1%	0%	0%	2%	0%	2%	0%	0%	0%	0%	0%	0%
Student says it is unlikely that matter existed anywhere else but the Big Bang.	0%	5%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says you can't make something out of nothing.	0%	3%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student says time began with the Big Bang.	0%	0%	19%	0%	0%	12%	0%	0%	17%	0%	0%	19%
Student says the Big Bang is the expansion of space.	0%	0%	4%	0%	18%	24%	0%	0%	0%	0%	0%	15%
Student says the Big Bang is the beginning of the universe.	1%	0%	0%	0%	0%	2%	3%	0%	0%	0%	0%	8%
Student says the Big Bang was an explosion.	2%	5%	11%	0%	9%	2%	7%	0%	17%	0%	0%	19%
Student says there was nothing before the Big Bang/everything began with the Big Bang.	2%	0%	15%	0%	0%	12%	0%	0%	0%	10%	0%	19%
Student says we don't/can't know.	1%	0%	4%	0%	0%	2%	0%	0%	0%	0%	0%	4%
Student gives irrelevant information.	0%	3%	0%	2%	0%	0%	3%	0%	0%	5%	0%	0%
Student gives some other reason not specified above.	0%	0%	11%	7%	9%	5%	3%	6%	50%	5%	0%	8%
Student has no idea.	0%	3%	4%	0%	0%	0%	0%	6%	0%	0%	0%	0%
Answer field is blank.	69%	53%	15%	54%	73%	7%	76%	67%	0%	45%	80%	0%

Table 8.20: Common reasoning elements used by students in their answers to Form B, Item 5: Independent of whether we know its true location, is there a center to the universe?

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says the universe is infinite/has no edges.	11%	17%	4%	22%
Student says the universe is the same everywhere.	1%	22%	2%	16%
Student says the Sun is at the center.	1%	0%	0%	0%
Student says everything has a center/must have a center.	11%	1%	9%	4%
Student says the center is where the Big Bang happened/where the universe began/where the universe is expanding from.	32%	9%	26%	4%
Student says things orbit the center of the universe.	3%	0%	6%	0%
Student says our ignorance prevents us from seeing the center.	11%	7%	19%	12%
Student says technological limitations prevent us from seeing the center.	1%	1%	2%	0%
Student says there is no center or the center changes because of expansion/objects are in motion.	11%	20%	26%	20%
Student says there is no center since the universe has no shape/an irregular shape/unknown shape.	0%	1%	2%	4%
Student gives irrelevant information.	1%	2%	2%	0%
Student gives some other reason not specified above.	10%	10%	9%	18%
Student has no idea.	4%	1%	9%	2%
Answer field is blank.	12%	19%	9%	16%

As in previous semesters, items 6 and 7 did not elicit any widespread reasoning elements beyond those used to assign students their overall numerical scores for these items. Table 8.15 shows that the change in the distribution of scores pre- to post-instruction was comparable for both populations. However, the change was slightly better for the non-lecture-tutorial students for both items, which is why they registered the larger normalized gains on both items.

8.4.3 Form C Responses

Table 8.21 shows the distribution of scores for the lecture-tutorial and non-lecture-tutorial populations pre- and post-instruction for the fall 2010. In general, the non-lecture-tutorial population shows a greater shift toward higher scores on Form C's items pre- to post-instruction. The one exception to this trend is item 2. These shifts in the distribution of scores are concordant with the normalized gains shown in Figure 8.1 above. In this section, we examine students' responses to Form C's items in more detail.

Table 8.21: Distribution of scores on Form C for fall 2010.

	LT Pre				Non-LT Pre			
	0	1	2	3	0	1	2	3
Item 1	0%	66%	16%	18%	5%	70%	9%	16%
Item 2	8%	61%	30%	1%	9%	58%	33%	0%
Item 3	15%	56%	5%	24%	14%	58%	9%	19%
Item 4	1%	73%	4%	22%	4%	79%	5%	12%
Item 5	1%	65%	10%	24%	5%	70%	12%	12%
Item 6	6%	35%	23%	35%	9%	39%	21%	32%
	LT Post				Non-LT Post			
	0	1	2	3	0	1	2	3
Item 1	1%	20%	24%	54%	0%	20%	34%	46%
Item 2	4%	50%	24%	21%	4%	61%	25%	11%
Item 3	4%	43%	10%	43%	13%	25%	5%	57%
Item 4	1%	56%	7%	36%	2%	57%	2%	39%
Item 5	1%	51%	7%	40%	5%	54%	5%	36%
Item 6	1%	47%	20%	31%	4%	36%	20%	41%

Figure 8.14 shows the percent of students who say the temperature of the universe has increased, decreased, stayed the same, and changed in response to item 1. The percent who say

decreased is less than 50% for both populations pre-instruction. Post-instruction, this response rises to around 80% for both groups. Figure 8.14 shows that the change in the percent of students saying "decreased" is greater for the non-lecture-tutorial than the lecture-tutorial students. This contributed to the larger normalized gain of the non-lecture-tutorial students on this item.

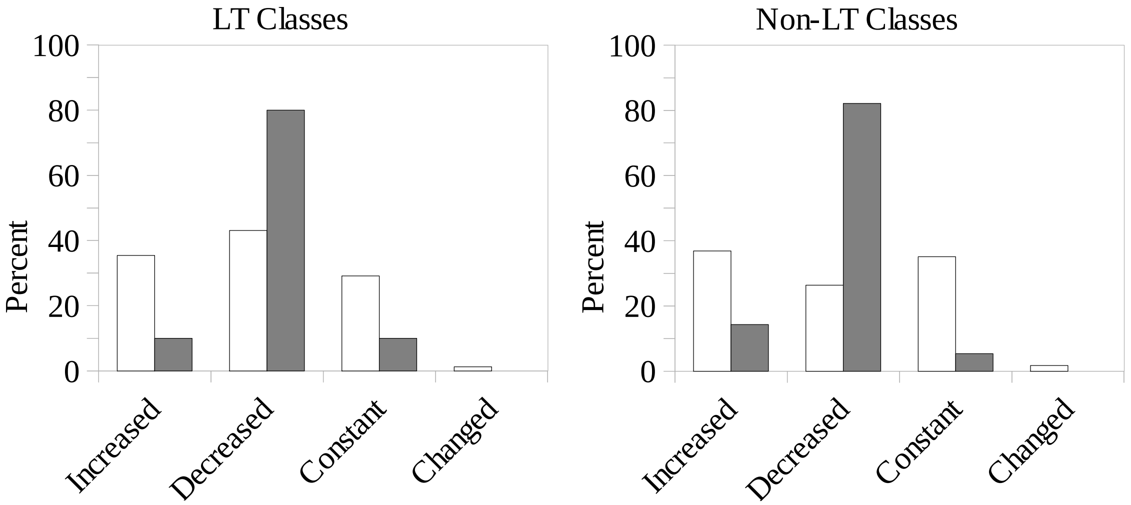


Figure 8.14: Percent of students who think the temperature of the universe has increased, decreased, stayed constant, or changed (without specifying if it went up or down). White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Table 8.22 gives the reasoning elements used by the lecture-tutorial students on item 1 and Table 8.23 gives the reasoning elements used by the non-lecture-tutorial students (as in previous chapters, I = the temperature increased, D = the temperature decreased, S = the temperature stayed the same, and C = the temperature changed). Among students who did not say the temperature decreased, the most commonly invoked reasons involve the births, deaths, and changes during the lives of stars and planets. This is consistent with results from previous semesters. Among students who did say the temperature decreased, the most common explanation was the expansion of the universe. 41% of lecture-tutorial students who said "decreased" pre-instruction talked about expansion in their answers. Post-instruction, this percent rose to 70%. This percent remained approximately constant for the non-lecture-tutorial students (60% pre-instruction and 59% post-instruction). This may indicate a positive effect of the lecture-tutorials, although we cannot say for certain since many students did not justify their answers.

Table 8.24 gives the reasoning elements used by students in their responses to item 2. Among the listed reasoning elements, we considered the first two to be correct. Compared to the non-lecture-tutorial students, a greater percent of lecture-tutorial students used these two reasoning elements post-instruction. Additionally, the percent of lecture-tutorial students claiming the amount of matter does not change dropped from 51% pre-instruction to 44% post-instruction. This percent actually rose for the non-lecture-tutorial population, from 44% pre-instruction to 47% post-instruction. Both groups exhibited a decrease in the percent of students claiming that the amount of matter increases as objects in the universe form and/or evolve. These response patterns help explain why the lecture-tutorial population achieved a higher learning gain on this item than the non-lecture-tutorial population.

Figure 8.15 shows the percent of students who answered “increased,” “decreased,” ”stayed the same,” and “changed” to item 3. While both groups show a shift toward “decreased” pre- to post-instruction, this shift is greatest for the non-lecture-tutorial population. This explains in part why the non-lecture-tutorial students had a larger normalized gain on item 3 than their lecture-tutorial counterparts.

Tables 8.25 and 8.26 show the reasoning elements used by the lecture-tutorial and non-lecture-tutorial populations, respectively. Many students who gave an incorrect answer talked about how objects form over time in the universe. This was also a popular answer in previous semesters’ data. Among students that said the density has decreased, a majority cited the expansion of the universe in their answers. For the lecture-tutorial students, 79% talked about expansion pre-instruction compared to 83% post-instruction. For the non-lecture-tutorial students, this percent went from 64% pre-instruction to 91% post-instruction. The fact that the non-lecture-tutorial population has a greater change in the percent of students talking about expansion also explains why they achieved the larger normalized gain.

Items 4-6 all explore the relationship between the expansion of the universe, lookback times, and the distances light travels. The non-lecture-tutorial population outperformed the lecture-

Table 8.22: Common reasoning elements used by lecture-tutorial students in their answers to Form C, Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	LT Pre				LT Post			
Reasoning Element	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	7%	41%	4%	0%	14%	70%	29%	0%
Student talks about the birth/formation of stars/a star.	18%	6%	22%	100%	29%	4%	43%	0%
Student talks about the birth/formation of planets/a planet.	0%	0%	4%	0%	14%	2%	0%	0%
Student talks about the birth/formation of galaxies/a galaxy.	0%	0%	4%	0%	14%	0%	0%	0%
Student talks about the birth/formation of unspecified objects.	4%	0%	0%	0%	14%	2%	0%	0%
Student talks about the death of stars.	7%	9%	22%	0%	14%	0%	14%	0%
Student talks about the death of planets.	0%	0%	4%	0%	0%	0%	0%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of stars/a star.	25%	6%	17%	0%	14%	2%	29%	0%
Student talks about changes during the lives of planets/a planet.	29%	15%	17%	0%	0%	2%	0%	0%
Student talks about changes during the lives of galaxies/a galaxy.	0%	0%	0%	0%	14%	0%	0%	0%
Student talks about changes during the lives of unspecified objects.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about how the universe is big.	0%	0%	4%	0%	0%	0%	0%	0%
Student talks about competing effects canceling out.	0%	0%	22%	0%	0%	0%	29%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	14%	12%	13%	0%	14%	4%	0%	0%
Student says s/he has no idea.	0%	3%	4%	0%	14%	4%	0%	0%
Response field left blank.	0%	0%	0%	0%	0%	0%	0%	0%

Table 8.23: Common reasoning elements used by non-lecture-tutorial students in their answers to Form C, Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Non-LT Pre				Non-LT Post			
Reasoning Element	I	D	S	C	I	D	S	C
Student talks about expansion and/or how the density of the universe changed.	24%	60%	0%	0%	13%	59%	0%	0%
Student talks about the birth/formation of stars/a star.	10%	7%	20%	0%	25%	2%	0%	0%
Student talks about the birth/formation of planets/a planet.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the birth/formation of galaxies/a galaxy.	0%	0%	0%	0%	13%	2%	0%	0%
Student talks about the birth/formation of unspecified objects.	0%	7%	5%	0%	0%	2%	33%	0%
Student talks about the death of stars.	5%	7%	25%	0%	13%	4%	0%	0%
Student talks about the death of planets.	0%	0%	5%	0%	0%	0%	0%	0%
Student talks about the death of galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the death of unspecified objects.	0%	0%	5%	0%	0%	0%	0%	0%
Student talks about changes during the lives of stars/a star.	0%	13%	5%	0%	25%	4%	0%	0%
Student talks about changes during the lives of planets/a planet.	38%	7%	0%	0%	13%	0%	0%	0%
Student talks about changes during the lives of galaxies/a galaxy.	0%	7%	0%	0%	0%	0%	0%	0%
Student talks about changes during the lives of unspecified objects.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about how the universe is big.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about competing effects canceling out.	0%	7%	15%	0%	0%	0%	0%	0%
Student gives irrelevant information.	5%	0%	5%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	10%	7%	45%	0%	0%	2%	0%	0%
Student says s/he has no idea.	14%	7%	5%	0%	0%	0%	0%	0%
Response field left blank.	0%	0%	0%	0%	0%	0%	0%	0%

Table 8.24: Common reasoning elements used by students in their answers to Form C, Item 2: How does the total amount of matter in the universe right now compare to the total amount of matter in the universe at the very beginning of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says there was no matter/only energy in the beginning.	0%	14%	0%	5%
Student talks about the temperature cooling and/or that matter formed from energy.	1%	14%	0%	11%
Student says there is more matter now because the universe is expanding.	4%	7%	0%	11%
Student says the amount of matter increases as objects form and/or evolve.	19%	10%	26%	19%
Student says the amount of matter decreases as objects form and/or evolve.	6%	4%	7%	4%
Student says the amount of matter increases as objects interact and/or die.	3%	3%	2%	0%
Student says the amount of matter decreases as objects interact and/or die.	3%	1%	2%	2%
Student says the amount of matter does not change.	51%	44%	44%	47%
Student gives irrelevant information.	0%	1%	2%	0%
Student gives other reason not specified above.	3%	6%	7%	8%
Student says s/he has no idea.	3%	0%	7%	0%
Response field left blank or no reason given.	14%	9%	7%	8%

Table 8.25: Common reasoning elements used by lecture-tutorial students in their answers to Form C, Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	LT Pre				LT Post			
Reasoning Element	I	D	S	C	I	D	S	C
Students says the amount of matter in the universe has changed.	0%	0%	0%	0%	6%	0%	0%	0%
Student says the amount of matter in the universe has not changed.	0%	21%	8%	0%	0%	8%	0%	0%
Student says the size of the universe has changed/stuff spread out.	7%	79%	0%	33%	6%	83%	8%	0%
Student says the size of the universe has not changed/stuff hasn't spread out.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time.	39%	0%	23%	33%	50%	3%	15%	0%
Student talks about objects evolving over time.	11%	4%	0%	33%	6%	0%	0%	0%
Student talks about objects dying/breaking up.	7%	8%	23%	0%	17%	3%	15%	0%
Student talks about the effects of gravity.	11%	0%	8%	0%	11%	0%	8%	0%
Student talks about matter changing forms.	0%	0%	0%	0%	0%	0%	0%	0%
Student gives some other reason not specified above.	7%	0%	23%	0%	6%	3%	8%	100%
Student gives irrelevant information.	0%	4%	0%	0%	0%	0%	0%	0%
Student says s/he has no idea.	7%	0%	0%	0%	0%	0%	0%	0%
Answer field is blank or no response given.	0%	0%	0%	0%	0%	0%	0%	0%

Table 8.26: Common reasoning elements used by non-lecture-tutorial students in their answers to Form C, Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

	Non-LT Pre				Non-LT Post			
Reasoning Element	I	D	S	C	I	D	S	C
Students says the amount of matter in the universe has changed.	13%	0%	0%	0%	0%	0%	0%	0%
Student says the amount of matter in the universe has not changed.	0%	21%	6%	0%	0%	9%	14%	0%
Student says the size of the universe has changed/stuff spread out.	13%	64%	0%	0%	0%	91%	14%	0%
Student says the size of the universe has not changed/stuff hasn't spread out.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about objects forming over time.	44%	0%	12%	0%	57%	0%	0%	0%
Student talks about objects evolving over time.	0%	14%	0%	0%	0%	0%	0%	0%
Student talks about objects dying/breaking up.	6%	0%	6%	0%	0%	3%	0%	0%
Student talks about the effects of gravity.	0%	0%	0%	0%	0%	0%	29%	0%
Student talks about matter changing forms.	0%	0%	0%	50%	0%	0%	14%	0%
Student gives some other reason not specified above.	0%	0%	24%	0%	14%	6%	14%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%
Student says s/he has no idea.	13%	0%	0%	0%	14%	3%	0%	0%
Answer field is blank or no response given.	0%	0%	0%	0%	0%	0%	0%	0%

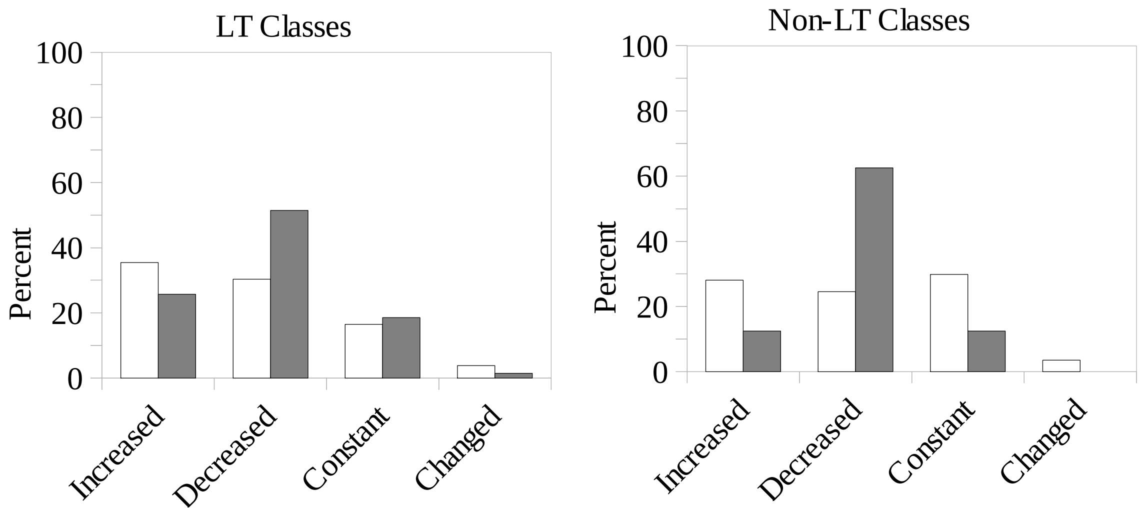


Figure 8.15: Percent of students who think the density of matter in the universe has increased, decreased, stayed constant, or changed (without specifying if it went up or down). White bars represent pre-instruction responses and grey bars represent post-instruction responses.

tutorial population on all three items. Figures 8.16-8.18 show the percent of students in each group who circled choices A, B, C, and D. In every case, the non-lecture-tutorial students had a greater, positive change in the percent circling the correct answer (A for item 4, C for item 5, and B for item 6) than the lecture-tutorial students. In fact, Figure 8.18 shows that the percent of lecture-tutorial students who circled B for item 6 actually decreased pre- to post-instruction. Why are the lecture-tutorial students not performing as well as the non-lecture-tutorial students?

Tables 8.27-8.32 show the reasoning elements used by both populations of students, pre- and post-instruction, in their answers to items 4-6. In general, we see the same categories of responses in the fall 2010 data as we did in the spring 2010 data. We can see why the non-lecture-tutorial students did better than the lecture-tutorial students when we examine the reasons students gave when they selected the correct answer. For example, on item 4 the percent of lecture-tutorial students who selected A and talked about the expansion of the universe did not change much pre- to post-instruction (85% pre-instruction compared to 83% post-instruction). For the non-lecture-tutorial students, the percent who chose A and talked about expansion went from 80% pre-instruction to 96% post-instruction. For another example, look at the responses to item 5 (Tables

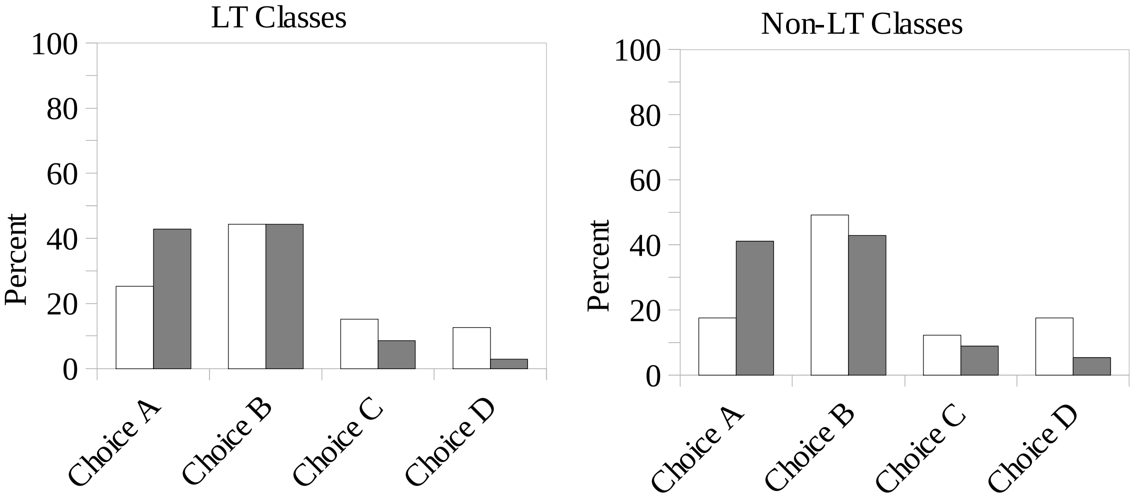


Figure 8.16: Percent of students who selected choice A, B, C, or D in response to item 4 on Form C. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

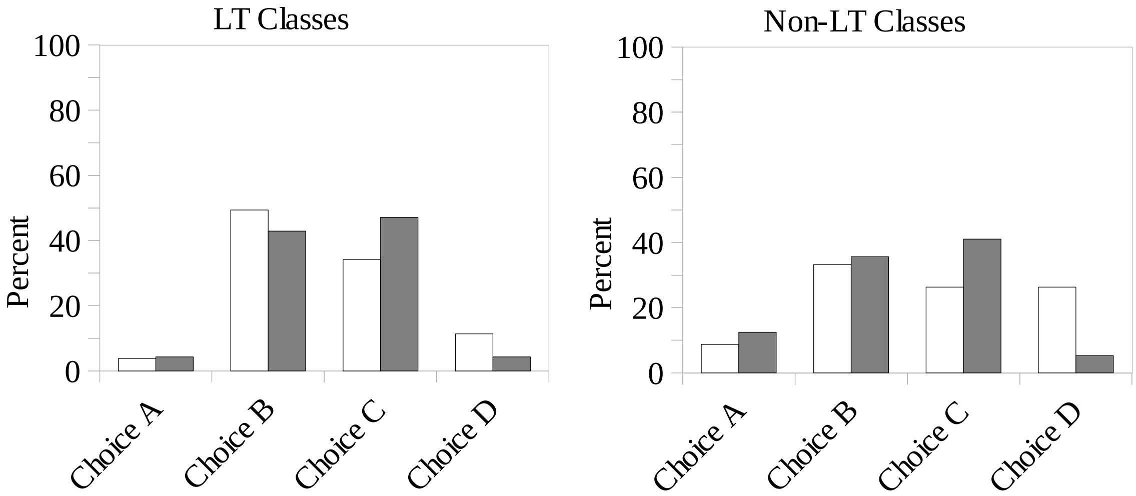


Figure 8.17: Percent of students who selected choice A, B, C, or D in response to item 5 on Form C. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

8.29 and 8.30). 70% of the lecture-tutorial students who chose C for item 5 also talked about the expansion of the universe pre-instruction. Post-instruction, this percent was 85%. Among the non-lecture-tutorial students, this percent went from 47% pre-instruction to 87% post-instruction. Finally, consider item 6. While the percent of students who reasoned that 13 billion years - 8 billion years = 5 billion years hovered around 60% pre- and post-instruction for both populations,

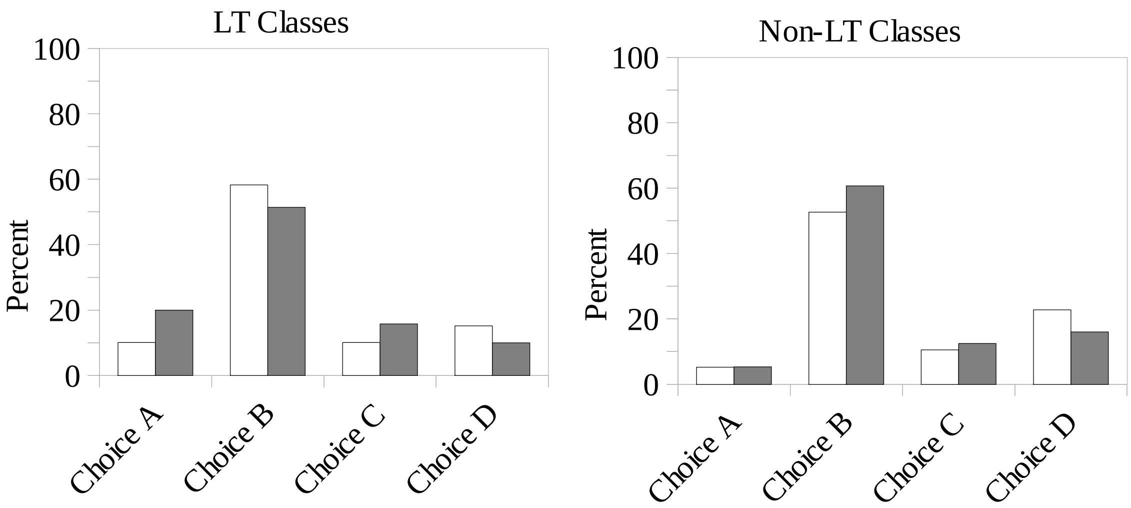


Figure 8.18: Percent of students who selected choice A, B, C, or D in response to item 6 on Form C. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

the change in the percent of students who talked about lookback time or the amount of time we need to see events happen in the universe was greater for the non-lecture-tutorial population. Overall, the non-lecture-tutorial students exhibited greater, positive changes in giving the right reasons for the right answers than did the lecture-tutorial students.

The responses to item 6 (Tables 8.31 and 8.32) also raise another issue. Pre- to post-instruction, there is a definite increase in the percent of students in both populations who chose an incorrect answer based on the expansion of the universe. This underscores Astro 101's students conceptual difficulties in thinking about lookback times in an expanding universe.

8.4.4 Form D Responses

Table 8.33 gives the distribution of scores on Form D for the lecture-tutorial and non-lecture-tutorial populations. The lecture-tutorial students exhibit larger improvements on all of Form D's items (pre- to post-instruction) than the non-lecture-tutorial students. This section examines both group's responses to these items.

The first two items of Form D ask students to identify the rotation curves for a solar system

Table 8.27: Common reasoning elements used by lecture-tutorial students in their answers to Form C, Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	LT Pre				LT Post			
Reasoning Element	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	85%	3%	17%	10%	83%	3%	67%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	63%	0%	0%	0%	71%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	8%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	40%	0%	0%	0%	50%
Student says we need more information on how galaxies moving together.	0%	0%	0%	40%	0%	0%	0%	50%
Student says 8 billion light-years is the minimum distance between Galaxies X and Y.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about time Galaxy X needs to see explosion.	10%	51%	17%	10%	37%	42%	50%	0%
Student says there's not enough information.	0%	0%	0%	80%	0%	0%	0%	50%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	10%	9%	25%	30%	3%	10%	0%	0%
Student says s/he has no idea.	0%	3%	8%	10%	0%	0%	0%	50%
Response field left blank or no reason given.	5%	20%	42%	0%	10%	13%	0%	0%

Table 8.28: Common reasoning elements used by non-lecture-tutorial students in their answers to Form C, Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Non-LT Pre				Non-LT Post			
Reasoning Element	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	80%	0%	14%	0%	96%	17%	40%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	46%	14%	0%	0%	46%	0%	0%
Student says light-years are shorter than normal years.	10%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	10%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	10%	0%	0%	0%	33%
Student says we need more information on how galaxies moving together.	0%	0%	14%	10%	0%	0%	0%	33%
Student says 8 billion light-years is the minimum distance between Galaxies X and Y.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about time Galaxy X needs to see explosion.	30%	36%	14%	0%	30%	63%	40%	33%
Student says there's not enough information.	0%	0%	14%	60%	0%	0%	0%	100%
Student gives irrelevant information.	0%	4%	0%	10%	0%	0%	0%	0%
Student gives other reason not specified above.	10%	7%	14%	30%	0%	13%	40%	67%
Student says s/he has no idea.	10%	4%	14%	30%	0%	0%	0%	0%
Response field left blank or no reason given.	0%	29%	43%	0%	4%	13%	20%	0%

Table 8.29: Common reasoning elements used by lecture-tutorial students in their answers to Form C, Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	LT Pre				LT Post			
Reasoning Element	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	33%	0%	70%	0%	67%	7%	85%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	46%	0%	0%	0%	60%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	4%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	22%	0%	0%	0%	67%
Student says we need more information on how galaxies moving together.	0%	0%	0%	22%	0%	0%	0%	67%
Student says we need more information on the rate of expansion.	0%	0%	0%	11%	0%	0%	0%	0%
Student says answer depends on where, exactly, events happen in Galaxies X and Y.	0%	0%	0%	0%	0%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	78%	0%	0%	0%	67%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	67%	13%	7%	33%	0%	20%	12%	0%
Student says s/he has no idea.	0%	0%	0%	11%	0%	0%	0%	0%
Response field left blank or no reason given.	0%	41%	19%	11%	33%	17%	6%	33%

Table 8.30: Common reasoning elements used by non-lecture-tutorial students in their answers to Form C, Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode? a) less than 8 billion light-years apart; b) exactly 8 billion light-years apart; c) more than 8 billion light-years apart; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Non-LT Pre				Non-LT Post			
Reasoning Element	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	20%	0%	47%	7%	57%	0%	87%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	47%	0%	0%	0%	60%	0%	0%
Student says light-years are shorter than normal years.	20%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	0%	0%	0%	0%	33%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	33%
Student says we need more information on the rate of expansion.	0%	0%	0%	0%	0%	0%	0%	0%
Student says answer depends on where, exactly, events happen in Galaxies X and Y.	0%	0%	7%	0%	0%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	40%	0%	0%	0%	33%
Student gives irrelevant information.	0%	0%	0%	7%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	16%	27%	33%	29%	15%	9%	0%
Student says s/he has no idea.	0%	5%	7%	13%	0%	0%	0%	0%
Response field left blank or no reason given.	60%	32%	13%	33%	14%	25%	9%	67%

Table 8.31: Common reasoning elements used by lecture-tutorial students in their answers to Form C, Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age? a) less than 5 billion years old; b) exactly 5 billion years old; c) more than 5 billion years old; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	LT Pre				LT Post			
Reasoning Element	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	25%	0%	0%	0%	50%	3%	45%	0%
Student talks about galaxies getting closer together.	0%	2%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	8%	0%	0%	0%	0%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	0%
Student reasons that 13-8=5/light traveled for 8 billion years.	0%	61%	0%	0%	7%	58%	18%	0%
Student talks about what Galaxy Y is like after the explosion.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about the star exploding more than 8 billion years ago.	0%	0%	13%	0%	0%	0%	0%	0%
Student talks about how it takes time to see events happen in the universe/lookback time.	25%	13%	0%	0%	36%	17%	18%	0%
Student says expansion affects age.	0%	0%	0%	0%	0%	0%	0%	0%
Student says answer depends on where, exactly, events happen in Galaxies X and Y.	0%	0%	0%	0%	0%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	25%	0%	0%	0%	25%
Student gives irrelevant information.	13%	0%	0%	0%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	4%	38%	8%	7%	6%	18%	50%
Student says s/he has no idea.	13%	0%	0%	25%	0%	0%	0%	0%
Response field left blank or no reason given.	25%	26%	50%	50%	29%	28%	18%	50%

Table 8.32: Common reasoning elements used by non-lecture-tutorial students in their answers to Form C, Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age? a) less than 5 billion years old; b) exactly 5 billion years old; c) more than 5 billion years old; d) there is not enough information to tell Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

	Non-LT Pre				Non-LT Post			
Reasoning Element	A	B	C	D	A	B	C	D
Student talks about the expanding universe/galaxies getting farther apart.	0%	0%	0%	0%	67%	0%	71%	0%
Student talks about galaxies getting closer together.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light needs 8 billion years to travel 8 billion light-years.	0%	3%	0%	0%	0%	0%	0%	0%
Student says light-years are shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are longer than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.	0%	0%	0%	0%	0%	0%	0%	0%
Student says light may be slowed by something between the two galaxies.	0%	0%	0%	0%	0%	0%	0%	0%
Student says we need more information on how galaxies moving apart.	0%	0%	0%	8%	0%	0%	0%	0%
Student says we need more information on how galaxies moving together.	0%	0%	0%	0%	0%	0%	0%	0%
Student reasons that 13-8=5/light traveled for 8 billion years.	0%	60%	0%	0%	0%	62%	0%	0%
Student talks about what Galaxy Y is like after the explosion.	0%	0%	17%	0%	0%	0%	0%	0%
Student talks about the star exploding more than 8 billion years ago.	0%	0%	0%	0%	0%	0%	0%	0%
Student talks about how it takes time to see events happen in the universe/lookback time.	33%	3%	0%	0%	0%	32%	0%	0%
Student says expansion affects age.	0%	0%	0%	0%	0%	0%	0%	0%
Student says answer depends on where, exactly, events happen in Galaxies X and Y.	0%	0%	0%	8%	0%	0%	0%	0%
Student says there's not enough information.	0%	0%	0%	31%	0%	0%	14%	44%
Student gives irrelevant information.	0%	0%	0%	8%	0%	0%	0%	0%
Student gives other reason not specified above.	0%	7%	17%	15%	0%	6%	0%	56%
Student says s/he has no idea.	33%	0%	17%	31%	0%	0%	0%	0%
Response field left blank or no reason given.	33%	30%	50%	23%	33%	21%	14%	44%

Table 8.33: Distribution of scores on Form D for fall 2010.

	LT Pre				Non-LT Pre			
	0	1	2	3	0	1	2	3
Item 1	7%	73%	20%	-	0%	84%	16%	-
Item 2	10%	75%	15%	-	0%	88%	12%	-
Item 3	9%	35%	56%	-	4%	23%	74%	-
Item 4	13%	42%	45%	-	7%	33%	60%	-
Item 5	8%	42%	23%	27%	0%	49%	23%	28%
Item 6	11%	83%	5%	2%	0%	91%	4%	5%
Item 7	19%	46%	28%	7%	4%	65%	32%	0%
	LT Post				Non-LT Post			
	0	1	2	3	0	1	2	3
Item 1	0%	62%	38%	-	0%	86%	14%	-
Item 2	0%	39%	61%	-	0%	67%	33%	-
Item 3	1%	19%	80%	-	4%	31%	65%	-
Item 4	1%	31%	68%	-	8%	45%	47%	-
Item 5	2%	21%	13%	64%	2%	57%	12%	29%
Item 6	1%	61%	2%	37%	2%	63%	8%	27%
Item 7	3%	19%	19%	59%	12%	29%	35%	24%

(item 1) and a spiral galaxy (item 2). Figure 8.19 shows students' graph choices for item 1 and Figure 8.20 shows students' graph choices for item 2. Pre-instruction, the responses of both groups are scattered among the various graph choices. Post-instruction, the lecture-tutorial students are more likely to select the correct graphs for both items (graph 3 for item 1 and graph 2 for item 2), although a significant minority of students still choose graph 5 for item 1. The non-lecture-tutorial students do not perform as well on these items post-instruction. For item 1, the percent of students selecting graph 3 actually decreases by a slight amount. For item 2, the percent of students choosing graph 2 is greater than it was pre-instruction but not nearly as high as it is for the lecture-tutorial students. These results imply that the lecture-tutorial students are more likely to remember the correct rotation curves for solar systems and spiral galaxies.

Item 3 asked students to rank the speeds of three planets in the solar system based on the rotation curve they selected in item 1. Table 8.33 shows that the lecture-tutorial students were much more successful on this item than the non-lecture-tutorial students post-instruction. Table 8.34 shows the reasoning elements used by students to justify their rankings. The percent of

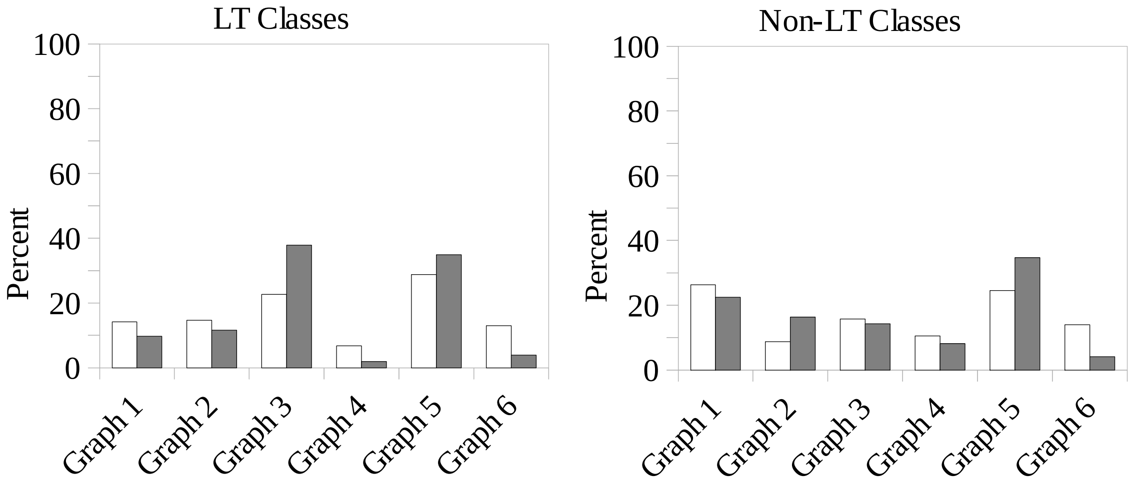


Figure 8.19: Students' graph choices for item 1 on Form D. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

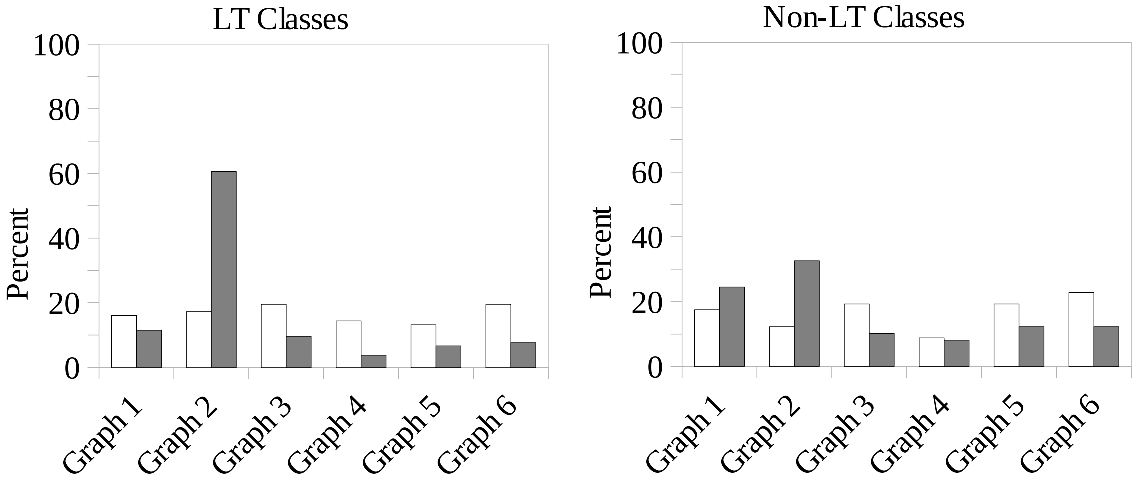


Figure 8.20: Students' graph choices for item 2 on Form D. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

lecture-tutorial students who said that the solar system's mass is concentrated in its center rose from 0% pre-instruction to 17% post-instruction, while for the non-lecture-tutorial students this percent remained constant at 2%. Pre-instruction, 14% of lecture-tutorial students say that closer planets move faster than farther planets because of the strength of the gravitational force they feel. Post-instruction, this goes up to 31%. Among the non-lecture-tutorial students, 25% make such a claim pre-instruction and 20% make this claim post-instruction. These responses further suggest that the “Dark Matter” lecture-tutorial has a positive effect on student learning.

Some of the reasoning elements listed in Table 8.34 suggest that some students may be activating various cognitive resources to justify their answers. For example, some students say that closer planets travel faster because they have a shorter distance to travel. Others say that closer planets move faster because they take less time to orbit than farther planets. Still others claim that the planets all take the same amount of time to orbit and thus closer planets travel at slower speeds than farther planets. Such responses may indicate that these Astro 101 students are using resources similar to those elucidated by Frank, Kanim, and Gomez (2008) in their study of resources students use to describe motion.

Item 4 is similar to item 3, except it asks students to rank the speeds of three stars based on the rotation curve they chose for the spiral galaxy in item 2. Again, Table 8.33 shows that the lecture-tutorial students did better than the non-lecture-tutorial students on this item post-instruction. Table 8.35 shows the reasoning elements used by students to justify their rankings. For both populations, the percent of students who explicitly claimed that stars orbiting a galaxy act like planets orbiting the Sun decreased pre- to post-instruction. However, the percent who explicitly stated that most of a galaxy’s mass is not in its center rose for the lecture-tutorial population from 1% to 13%. For the non-lecture-tutorial population, it went from 2% to 0%. Similarly, the percent of lecture-tutorial students who offered no justification for their rankings dropped from 45% to 18%, while for the non-lecture-tutorial population it rose from 16% to 31%. These patterns help explain why the lecture-tutorial students did better on item 4 than the non-lecture-tutorial students.

Figure 8.21 shows the percent of students circling each answer choice in item 5, which asks about the distribution of matter in solar systems. The most common choice, A (most of the mass in in the Sun), is also the correct choice. Note that the percent of lecture-tutorial students choosing A increases by almost 30% pre- to post-instruction, whereas it drops by 10% for the non-lecture-tutorial students. Table 8.36 shows the reasoning elements used by students to justify their answers. Note that one of the most common reasons students give for selecting either B or C is that they think the non-Sun objects in the solar system make a significant contribution to its overall

Table 8.34: Common reasoning elements used by students in their answers to Form D, Item 3: Rank the speeds at which Planets A, B, and C orbit the Sun. Explain your reason for ranking this way.

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says most of the solar system's mass is at its center.	0%	17%	2%	2%
Student says most of the solar system's mass is not in its center.	0%	0%	0%	0%
Student says planets travel at the same speed because they all take the same amount of time to orbit.	1%	0%	0%	0%
Student says planets travel at the same speed regardless of distance (no further explanation given).	2%	5%	4%	4%
Student says closer (farther) planets move faster (slower); no further explanation given.	18%	23%	7%	4%
Student says closer (farther) planets move faster (slower) because they have a shorter (longer) distance to travel.	11%	6%	19%	16%
Student says closer (farther) planets move faster (slower) because they need less (more) time to orbit.	4%	1%	5%	2%
Student says closer (farther) planets move slower (faster) because all planets take the same amount of time to orbit.	4%	0%	11%	10%
Student says closer (farther) planets move faster (slower) because they feel more (less) gravity.	14%	31%	25%	20%
Student says closer (farther) planets move faster (slower) because they feel more (less) centripetal force.	1%	0%	0%	0%
Student references the effects of gravity, but her response doesn't fit into the above categories.	2%	5%	4%	2%
Student gives irrelevant information.	2%	0%	2%	2%
Student give other reason not specified above.	10%	5%	7%	16%
Student says s/he has no idea.	1%	0%	0%	0%
Response field left blank or no reason given.	35%	14%	21%	27%

Table 8.35: Common reasoning elements used by students in their answers to Form D, Item 4: Rank the speeds at which Stars A, B, and C orbit the galaxy. Explain your reason for ranking this way.

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says stars orbiting the galaxy act like planets orbiting the Sun/answer is the same as item 3.	13%	2%	21%	4%
Student says most of the galaxy's mass is in its center.	0%	1%	0%	2%
Student says most of the galaxy's mass is not in its center.	1%	13%	2%	0%
Student says most of the galaxy's mass is in the spiral arms.	0%	3%	0%	0%
Student says most of the galaxy's mass is not in the spiral arms.	0%	0%	0%	0%
Student says most of the galaxy's mass is located outside/at the edges of the galaxy.	0%	1%	0%	2%
Student says most of the galaxy's mass is not located outside/at the edges of the galaxy.	0%	1%	0%	0%
Student says stars travel at the same speed because they all take the same amount of time to orbit.	1%	0%	0%	0%
Student says planets travel at the same speed regardless of distance because of dark matter.	1%	14%	0%	6%
Student says stars travel at the same speed regardless of distance (no further explanation given).	10%	13%	7%	18%
Student says closer (farther) stars move faster (slower); no further explanation given.	8%	3%	5%	2%
Student says closer (farther) stars move faster (slower) because they have a shorter (longer) distance to travel.	4%	3%	9%	2%
Student says closer (farther) stars move faster (slower) because they need less (more) time to orbit.	2%	1%	2%	0%
Student says closer (farther) stars move slower (faster) because all planets take the same amount of time to orbit.	4%	1%	9%	8%
Student says closer (farther) stars move faster (slower) because they feel more (less) gravity.	8%	2%	7%	2%
Student says the gravity of the galaxy is not concentrated in the galaxy's center due to dark matter.	0%	3%	0%	0%
Student says closer (farther) stars move faster (slower) because they feel more (less) centripetal force.	1%	0%	0%	0%
Student says stars farther out in the galaxy are affected more by dark matter than stars closer to the center.	0%	7%	0%	2%
Student references the effects of gravity, but her answer doesn't fit into the above categories.	3%	3%	2%	2%
Student gives irrelevant information.	2%	0%	2%	4%
Student give other reason not specified above.	9%	16%	32%	12%
Student says s/he has no idea.	6%	3%	12%	6%
Response field left blank or no reason given.	45%	18%	16%	31%

mass. These results are consistent with the idea that the "Dark Matter" lecture-tutorial helps move students toward the idea that the Sun accounts for the vast majority of the solar system's mass.

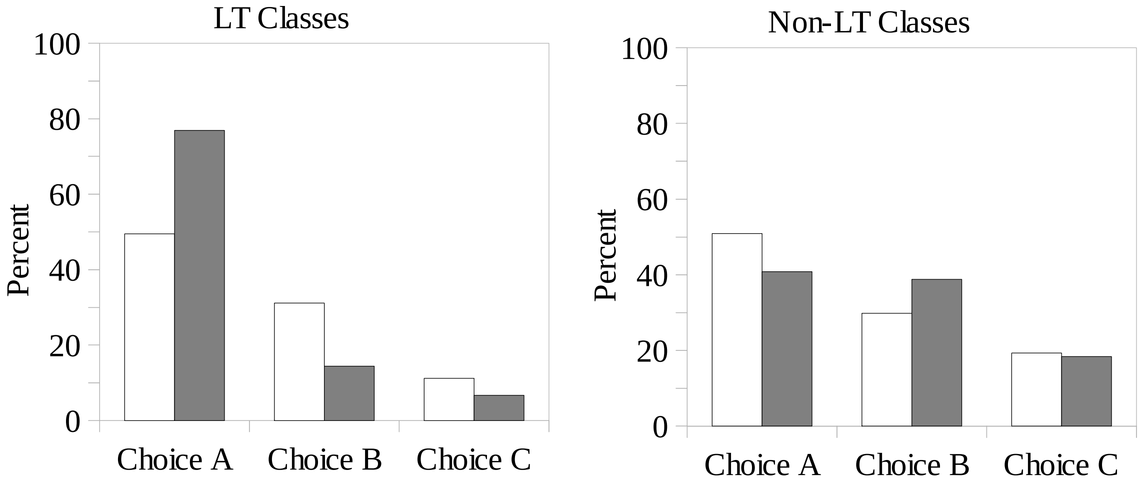


Figure 8.21: Percent of students who selected choice A, B, or C in response to item 5 on Form D. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

Item 6 is similar to item 5, except it asks about the distribution of mass in a spiral galaxy. The correct answer to item 6 is C. C is never chosen by a majority of students in either population; that honor belongs to A, suggesting that many students think the mass of a galaxy is concentrated in its center like the mass of a solar system. This is consistent with the response patterns shown in Table 8.37. Nevertheless, more students select choice C after instruction. The change in the percent of lecture-tutorial students choosing C is greater than the change for the non-lecture-tutorial students. Furthermore, Table 8.37 shows that the lecture-tutorial students are much more likely to discuss dark matter in their answers post-instruction than the non-lecture-tutorial students. The fact that they are more likely to mention dark matter is significant because Form D's items never use the words "dark matter."

Item 7 asks students to compare planets orbiting the Sun to stars orbiting the galaxy. Table 8.38 shows that, pre- to post-instruction, 49% more lecture-tutorial students explicitly state that planets do not act like stars. This difference is 28% for the non-lecture-tutorial students. Furthermore, the lecture-tutorial students are more likely to claim that the distribution of matter in

Table 8.36: Common reasoning elements used by lecture-tutorial students in their answers to Form D, Item 5: Based on your previous answers, how is matter distributed in solar systems? Pick the best answer from the following choices (a-c). a) Most of the matter in the solar system is located in the Sun. b) Most of the matter in the solar system is evenly distributed throughout the Sun and planets. c) My thinking is different than a and b. Explain your reasoning.

	LT Pre			LT Post			Non-LT Pre			Non-LT Post		
Reasoning Element	A	B	C	A	B	C	A	B	C	A	B	C
Student says the Sun is big/the most massive object in the Solar System.	45%	0%	5%	64%	7%	0%	59%	0%	0%	60%	0%	0%
Student says the Sun has a large gravitational force.	10%	0%	0%	8%	0%	0%	14%	0%	0%	0%	0%	0%
Student says the Solar System's gravity/mass is concentrated at the location around which everything orbits.	5%	0%	0%	4%	0%	0%	10%	0%	0%	0%	0%	0%
Student says other non-Sun objects in the Solar System contribute to the mas of the Solar System.	0%	18%	27%	0%	13%	29%	0%	41%	18%	5%	21%	33%
Student says the mass of all the non-Sun objects in the Solar System exceeds the Sun's mass.	0%	2%	9%	0%	7%	0%	0%	6%	9%	0%	0%	0%
Student says most of the mass of the Solar System is at its edge.	0%	0%	0%	0%	0%	14%	0%	0%	0%	0%	0%	0%
Student gives irrelevant information.	2%	0%	0%	1%	0%	0%	0%	0%	0%	0%	0%	0%
Student give other reason not specified above.	2%	15%	32%	13%	13%	43%	3%	12%	64%	10%	26%	67%
Student says s/he has no idea.	0%	3%	0%	0%	0%	0%	0%	0%	0%	5%	0%	0%
Response field left blank or no reason given.	39%	64%	32%	18%	60%	14%	31%	47%	9%	25%	58%	0%

Table 8.37: Common reasoning elements used by lecture-tutorial students in their answers to Form D, Item 6: Based on your previous answers, how is matter distributed in spiral galaxies? Pick the best answer from the following choices (a-c). a) Most of the matter in the galaxy is located in the center. b) Most of the matter in the galaxy is located in the spiral arms. c) My thinking is different than a and b. Explain your reasoning.

	LT Pre			LT Post			Non-LT Pre			Non-LT Post		
Reasoning Element	A	B	C	A	B	C	A	B	C	A	B	C
Student says the distribution of mass in the galaxy is like the distribution of mass in the Solar System/the answer is the same as item 5.	6%	2%	8%	4%	0%	0%	8%	0%	20%	0%	0%	0%
Student says most of the mass is at the center because that's where the density is highest/galaxy is brightest.	22%	2%	8%	26%	0%	0%	26%	0%	0%	20%	0%	0%
Student says there is little matter in the center.	0%	2%	0%	0%	3%	0%	0%	7%	0%	0%	0%	0%
Student says most of the matter in the galaxy is at its edge/outskirts.	0%	0%	0%	0%	3%	22%	0%	0%	0%	0%	0%	18%
Student says most of the matter in the galaxy is in its stars.	5%	6%	8%	9%	13%	2%	5%	0%	0%	7%	13%	6%
Student says matter is evenly distributed throughout the galaxy.	0%	2%	31%	0%	8%	39%	0%	7%	60%	0%	6%	53%
Student says most of the mass of the galaxy is dark matter.	1%	0%	0%	0%	3%	15%	0%	0%	20%	0%	6%	24%
Student says there is more dark matter in the spiral arms/disk than anywhere else.	0%	0%	0%	0%	20%	2%	0%	0%	0%	0%	6%	0%
Student says dark matter is located outside the parts of the galaxy you can see.	0%	0%	0%	0%	3%	24%	0%	0%	0%	0%	0%	12%
Student says matter is brought closer together as it moves toward the center.	5%	0%	0%	4%	0%	0%	8%	0%	0%	13%	0%	0%
Student talks about the black hole in the galaxy's center.	1%	0%	0%	4%	0%	0%	0%	0%	0%	0%	0%	0%
Student says most of the galaxy's gravity is located in its spiral arms.	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%
Student says gravity is evenly distributed throughout the galaxy.	0%	0%	0%	0%	3%	0%	0%	0%	0%	0%	0%	0%
Student says the galaxy's gravity/mass is concentrated at the location around which everything orbits.	10%	0%	0%	4%	0%	0%	18%	0%	0%	0%	0%	0%
Student references dark matter, but her answer doesn't fit into any of the above categories.	0%	4%	8%	0%	0%	10%	0%	0%	0%	0%	0%	0%
Student gives irrelevant information.	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%	0%
Student give other reason not specified above.	4%	13%	0%	9%	13%	5%	8%	36%	0%	7%	44%	6%
Student says s/he has no idea.	4%	8%	23%	0%	0%	0%	3%	7%	0%	0%	0%	0%
Response field left blank or no reason given.	49%	65%	23%	43%	40%	2%	45%	43%	0%	60%	25%	12%

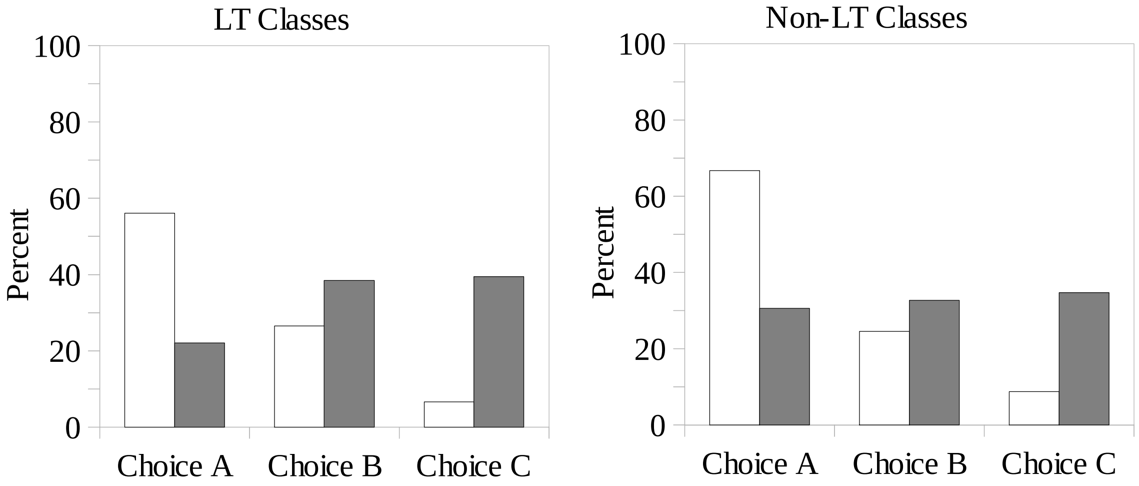


Figure 8.22: Percent of students who selected choice A, B, or C in response to item 6 on Form D. White bars represent pre-instruction responses and grey bars represent post-instruction responses.

galaxies is different than the distribution of matter in solar systems, and they are also more likely to say that the velocities of stars/rotation curves of galaxies are different than the velocities of planets/rotation curves of solar systems. These results strongly support the claims for the efficacy of the "Dark Matter" lecture-tutorial.

8.5 Validity

In the previous two chapters, I discussed the various pieces of evidence in support of the validity of our interpretations of students' scores on Forms A-D. In this section, I return one final time to some of the issues raised in Chapters 6 and 7. I specifically want to address three of the assumptions listed in Chapter 6: 1) The students who take the surveys are representative of the target population of Astro 101 students – that is, we can generalize our results; 2) Astro 101 students correctly read and interpret our survey items; and 3) Differences in the learning gains of students who have and have not used the lecture-tutorials are due to the lecture-tutorials and not some other variable.

First, are our results generalizable to the entire population of Astro 101 students at American colleges and universities? Our data from previous semesters was drawn from classes taught at just

Table 8.38: Common reasoning elements used by students in their answers to Form D, Item 7: Based on your answers for Questions 1-6, do stars orbiting the center of a galaxy act like planets orbiting the Sun? If yes, explain why. If no, explain why not.

Reasoning Element	LT Pre	LT Post	Non-LT Pre	Non-LT Post
Student says orbiting stars act like orbiting planets.	47%	20%	65%	29%
Student says orbiting stars do not act like orbiting planets.	25%	74%	21%	49%
Student says orbiting stars do not act like orbiting planets, but only because of superficial differences.	7%	3%	11%	8%
Student says stars move slower than planets.	2%	0%	0%	0%
Student says stars have irregular orbits/don't orbit a specific point, unlike planets.	8%	5%	12%	10%
Student says the distribution of matter in galaxies is different from the distribution of matter in solar systems.	2%	35%	2%	10%
Student says the velocities of stars/rotation curves of galaxies are different than the velocities of planets/rotation curves of solar systems.	8%	36%	2%	14%
Student talks about the black hole at the center of the galaxy.	2%	2%	2%	0%
Student references the effects of gravity, but her responses does not fit into any of the previous categories.	10%	12%	21%	10%
Student references the effects of dark matter, but her response does not fit into any of the previous categories.	0%	2%	2%	0%
Student says her previous responses were the same for the solar system and the galaxy.	1%	0%	2%	2%
Student says her previous responses were not the same for the solar system and the galaxy.	0%	2%	0%	0%
Student gives irrelevant information.	0%	1%	2%	0%
Student give other reason not specified above.	7%	7%	11%	12%
Student says s/he has no idea.	6%	0%	5%	6%
Response field left blank or no reason given.	17%	2%	0%	6%

three institutions: the University of Colorado at Boulder, the University of Arizona, and Syracuse University. The situation is very different for our fall 2010 data. As described in Section 8.1 above, our fall 2010 data is taken from fourteen classes taught at eleven different institutions. These span a range of class size and institutional type. Our sample includes classes of 100 or more students, classes with fewer than ten students, and classes with enrollments in between these numbers. It includes community college classes, classes taught at liberal arts colleges, and classes taught at large research-focused institutions. We have both public and private colleges in our sample as well. When we look at the data we collected for the entire study, we have surveyed a total of 2318 students pre-instruction and 2041 students post-instruction. 1709 of the pre-instruction responses and 1527 of the post-instruction responses came from students in lecture-tutorial classes. 609 of the pre-instruction responses and 514 of the post-instruction responses came from students in non-lecture-tutorial classes. The fact that these classes continually reveal the same set of difficulties with cosmology suggests that the results of this study apply to the broader population of American Astro 101 students.

Second, do Astro 101 students correctly read and interpret our survey items? In previous semesters, we detected a handful of items that did not function as we intended. This led us to make revisions to our surveys after the fall 2009 and spring 2010 semesters. Do we detect any issues with the fall 2010 versions of our surveys? We did not notice any patterns in students' written responses indicative of problematic items. Furthermore, I interviewed nine Astro 101 students during the fall 2010 semesters using the fall 2010 surveys and following the same procedure for the interviews I conducted in previous semesters (see Section 6.6 for more details). Table 8.39 lists each interviewed student by his/her pseudonym as well as the forms he/she reviewed and the order in which they were reviewed. Interviewed students in previous semesters helped highlight everything from minor typos to items that caused consistent and persistent confusions. However, none of the nine interviewed students in the fall of 2010 found any problems with any of the items on any of the surveys. This suggests that our previous revisions have produced surveys that comprehensible in their entireties to Astro 101 students.

Table 8.39: The forms each student responded to during his/her interview in the fall 2010. The numbers denote the order in which I presented the surveys to the student.

Student	Form A	Form B	Form C	Form D
Molly	2	-	3	1
Patrick	3	2	-	1
Eduardo	2	-	1	3
Stan	3	-	1	2
Vanessa	1	3	-	2
Cecelia	-	1	3	2
Timothy	-	1	2	3
Tucker	1	2	3	-
Brett	1	3	2	-

Finally, to what extent might other variables explain any differences between the results of the lecture-tutorial and non-lecture-tutorial populations? During the spring of 2010 I observed two Astro 101 classes at the University of Colorado at Boulder and found they did not differ in terms of the amount of time spent covering cosmology, despite the fact that one class used the lecture-tutorials while the other did not (see Section 7.5). I repeated these observations in the fall of 2010 for Class H and Class Q, both of which were taught at the University of Colorado at Boulder. Neither class had recitation or laboratory sections. Class H used the cosmology lecture-tutorials while Class Q did not. Did Class H end up spending more time on cosmology than Class Q?

According to my observations, the answer is no. I observed 76% of Class H's lectures and 85% of Class Q's lectures, but none of the lectures I missed covered cosmology. Class H spent 261 minutes out of the 1440 minutes I observed (or 18% of the time) covering cosmology, including the lecture-tutorials. Class Q spend 418 minutes out of the 1425 minutes I observed (or 29% of the time) on cosmology. Class H did not spend more time on cosmology even though it used the lecture-tutorials.

Class H did differ from Class Q in the amount of time spent on interactive engagement. 15% of the time Class H spent on cosmology was devoted to think-pair-share questions. The lecture-tutorials took up another 39%. Thus, the Class H spent 55% of its time on cosmology using some kind of interactive engagement technique. In contrast, the only interactive engagement technique

Class Q used was think-pair-share questions. These accounted for 19% of the time Class Q spent on cosmology. The lecture-tutorials thus mark the largest difference in how Class H and Class Q used their time on cosmology.

Overall, the data discussed in this section strengthen the validity argument we have made throughout Chapters 8 and 7. We have evidence that our results can be generalized to the broader Astro 101 population, we have evidence that students do interpret our survey items as we intended, and we have evidence that use of the lecture-tutorials comprises the most significant difference between the two Astro 101 classes I observed. This evidence, combined with the validity evidence described in Chapters 8 and 7, strengthens our confidence in our conclusions.

8.6 Summary of Fall 2010 Results

This chapter presented our results from the fall 2010 data. Overall, this data reveals many of the same conceptual and reasoning difficulties we uncovered in the fall 2009 and spring 2010 data. We also obtained a better understanding of students' difficulties with dark matter using our revised and expanded version of Form D. The data presented in this chapter suggests that this revised Form D is more reliable than in previous semesters. Students' responses to this form also suggest that many students may be using cognitive resources uncovered by previous studies (e.g. Frank, Kanim, and Gomez 2008) when constructing their responses to items on Form D. Finally, we presented additional evidence for the validity of our survey score interpretations. This evidence strengthens our ongoing validity argument and lends support to the generalizability and trustworthiness of our results.

Our comparison of the lecture-tutorial and non-lecture-tutorial populations yielded a couple interesting results. First, the lecture-tutorial students showed larger learning gains on Forms A and D than their non-lecture-tutorial counterparts. The results for Form D are consistent with previous semesters' results. The results from Form A are interesting because previous semesters' data failed to show any difference in the gains of lecture-tutorial and non-lecture-tutorial students. Consequently, both Form A and the "Hubble's Law" lecture-tutorial went through several revisions.

The data from Form A for this semester suggest that these revisions may have achieved some success.

On the other hand, the fall 2010 data do not show larger gains for the lecture-tutorial students compared to the non-lecture-tutorial students for Forms B and C. This result was surprising, since the lecture-tutorial students in previous semesters demonstrated significantly larger gains on Forms B and C than the non-lecture-tutorial students. What might cause such a discrepancy between the fall 2010 data and the data from the prior two semesters?

One possibility is that the lecture-tutorials are having a positive effect, but our scoring procedures might not be detecting this effect. This could be the case for some items, such as item 2 on Form B as discussed above. However, the data described in Section 8.4 seems to rule this out as a viable possibility for the majority of items on Forms B and C. Additionally, we used the same scoring rubrics for many of these items in the fall 2010 as we did in previous semesters. The explanation seems to lie elsewhere.

There is one significant difference between the lecture-tutorial population in the fall 2010 and the lecture-tutorial populations of previous semesters: Fall 2010 was the first semester in which instructors from outside of our research group used the lecture-tutorials. This suggests that the differences in the results may have something to do with how different instructors implemented the lecture-tutorials. However, we must also consider the possibility that some of non-lecture-tutorial instructors in this study are engaging in instructional practices that are at least as effective as the cosmology lecture-tutorials in promoting students' conceptual cosmology knowledge. These possibilities are intriguing subjects for future studies, as described in the next chapter.

Chapter 9

Conclusions

This dissertation contains the findings of one of the first dedicated and systematic studies of Astro 101 students' difficulties with cosmology. It also describes the development and testing of a new suite of five cosmology lecture-tutorials. We designed these lecture-tutorials to help students overcome the most common difficulties revealed by this study. Section 9.1 provides an overview of these difficulties and the efficacy of the new lecture-tutorials. I then end this dissertation in Section 9.2 with a discussion of how this work might influence future AER studies.

9.1 Summary

For this study, we created four conceptual cosmology surveys, each of which focused on a different aspect of cosmology. Form A looked at students' difficulties in interpreting Hubble plots. Form B probed students' conceptualizations of the Big Bang and the expansion of the universe. Form C looked at students' ideas about the evolution of various properties of the universe over time. Form D looked at whether or not students understand why flat rotation curves for spiral galaxies are evidence for dark matter. What common conceptual and reasoning errors did these surveys uncover?

On Form A, we found that Astro 101 students experience widespread difficulties in interpreting Hubble plots. The most persistent of these difficulties occurred when we asked students to use Hubble plots to reason about the expansion rate of the universe. Many students struggled to find the correct Hubble plots for universes expanding at a constant rate, contracting at a constant rate, expanding at a faster rate over time, and expanding at a slower rate over time. Students did somewhat better at relating the expansion rate to the age of the universe, provided we do not use terms such as "older universe" and "younger universe," both of which are frequently misinterpreted by students. These results should caution Astro 101 instructors who use Hubble plots to talk about the expansion of the universe.

Students' responses to Form B support the claims of several previous papers (Lightman, Miller, and Leadbeater 1987; Lightman and Miller 1989; Lineweaver and Davis 2005; Prather, Slater, and Offerdahl 2003; Simonelli and Pilachowski 2004) on people's ideas about the Big Bang and the expansion of the universe. We encountered numerous instances of students who claim that the Big Bang was an explosion of pre-existing matter into empty space. Students who hold this model tend to think of the expansion of the universe refers to matter flying away from a central location and into previously unoccupied regions of space. We also found that a significant minority of students reject the idea that the universe is actually expanding; instead, they consider "expansion" a metaphor for the increase in our knowledge about the universe over time and/or for the formation of new objects in the universe. Some students also think that the Big Bang refers to the beginning of something smaller than the universe (such as the planet Earth), while others think that it was an event that happened to something smaller than the universe (such as an asteroid striking the Earth and killing off the dinosaurs). These findings reveal that many students have non-expert-like understandings of the Big Bang and the expansion of the universe.

We observed a wide variety of responses to items on Form C. In many cases, students appear to be extending or generalizing their intuitions and experiences to help them answer questions about how properties of the universe, such as its temperature and density, have changed over time. For example, some students incorrectly claim that the temperature of the universe must be increasing due to warming trends observed on Earth. On items addressing the relationship between distances, lookback times, and the expansion of the universe, we note two types of difficulties. In one type, students simply ignore the effects of expansion and answer these items as if the universe was static. In the other, students try to account for the effects of expansion, but their reasoning leads them in the wrong direction (e.g. they say a distance should have been smaller when it should have been larger). These results reveal the kinds of difficulties students face when they attempt to understand the evolution of the universe.

Finally, students' responses to Form D reveal that many of them struggle to understand why flat rotation curves are evidence for dark matter in spiral galaxies. In the fall of 2010, we saw several examples of students who appear to use some of the motion resources uncovered by Frank, Kanim, and Gomez (2008) to understand the orbits of planets in a Solar System and stars in a spiral galaxy. We also observed, over the course of the study, responses that suggest many students assume that the mass of a spiral galaxy is concentrated in its center, much as the mass of the Solar System is concentrated in the Sun. These difficulties persist post-instruction in the absence of any intervention.

What effect did our intervention, the cosmology lecture-tutorials, have on students? Figure 9.1 shows the normalized gains for the lecture-tutorial and non-lecture-tutorial populations for each of the three semesters we collected data. As Figure 9.1 shows, there are many cases in which the lecture-tutorial students achieve higher learning gains than the non-lecture-tutorial students. This suggests that the cosmology lecture-tutorials can help students learn more cosmology than traditional forms of instruction.

However, we must temper these conclusions with two observations. First, our sample of Astro 101 students always had normalized gains on Form A around 0.10 or smaller, regardless of whether or not they used the lecture-tutorials. Even though the difference in gains between the lecture-tutorial and non-lecture-tutorial students was statistically significant in the fall of 2010, we cannot claim any great victory in helping students learn to interpret Hubble plots. At the very least, these results suggest Astro 101 students are unlikely to select the correct Hubble plot and provide explanations for their selections that satisfy our stringent requirements for the maximum score, even after completing the lecture-tutorials.

Second, there are examples on Forms B and C of non-lecture-tutorial populations achieving higher normalized gains than lecture-tutorial populations. This is especially evident in the fall 2010

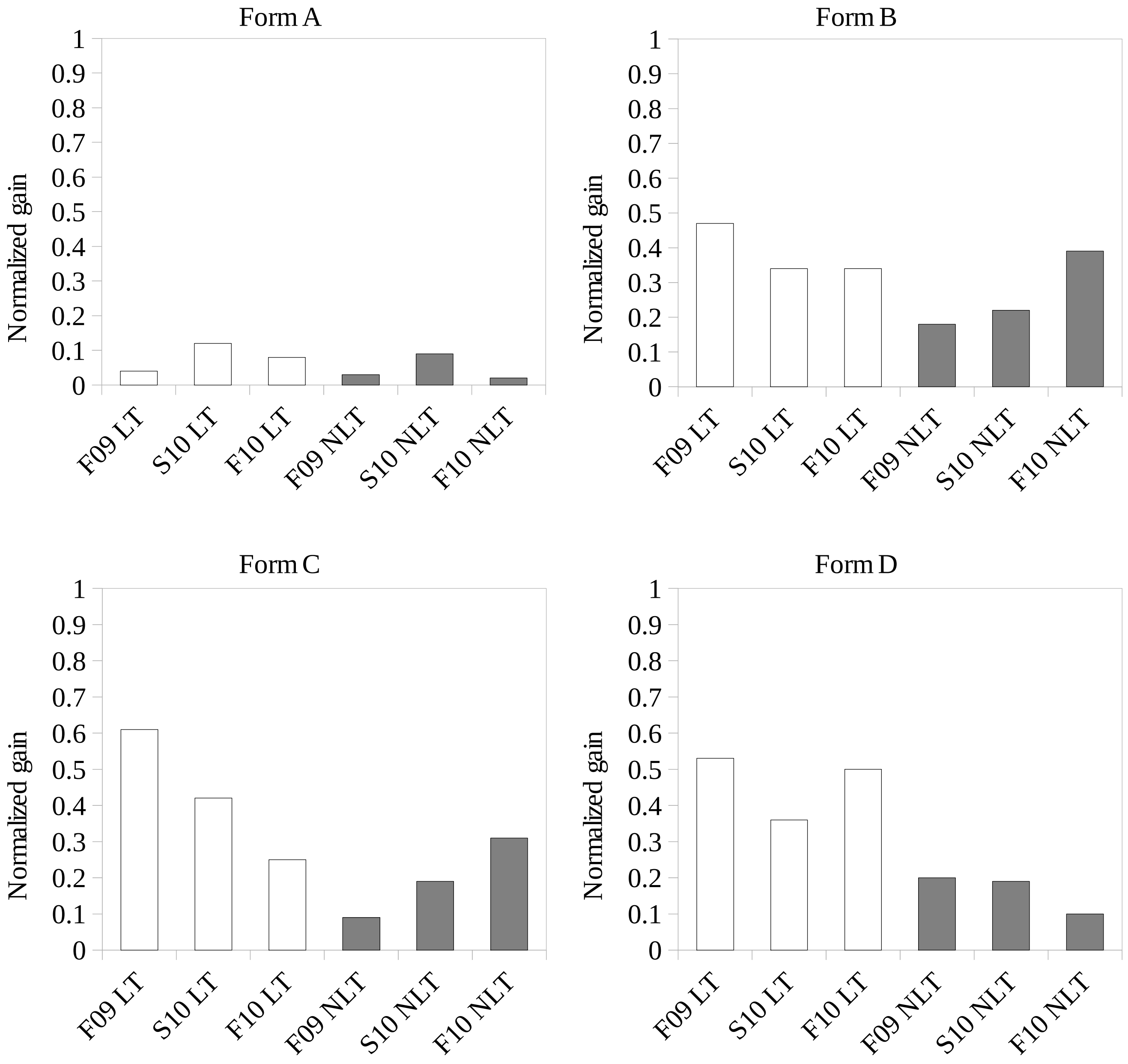


Figure 9.1: The normalized gains for Forms A-D for the fall 2009, spring 2010, and fall 2010. White bars correspond to lecture-tutorial groups and grey bars correspond to the non-lecture-tutorial groups.

data. How can we explain these results?

Perhaps a more fundamental question is "What causes the observed variation in gain scores?" Prior studies comparing the efficacy of interactive engagement to tradition instruction have investigated several variables (Hake 1998a; Hake 1998b; Prather *et al.* 2009; Rudolph *et al.* 2010). These include

(1) institution type,

(2) class size,

(3) time spent on interactive engagement,

(4) students' sex,

(5) students' ethnicity,

(6) students' prior math and science background,

(7) students' GPA,

(8) students' primary language,

(9) statistical fluctuations/random errors,

(10) systematic errors (which Hake 1998a further subdivides into question ambiguities and false positives, teaching to the test and test-question leakage, time on task, students' motivation, and Hawthorne/John Henry effects), and

(11) implementation practices.

These studies all provide evidence that the most important of these factors is instructors' implementation practices; all others are either irrelevant or of secondary importance (Hake 1998a; Hake 1998b; Prather *et al.* 2009; Rudolph *et al.* 2010). We therefore infer that the variation in observed normalized gains on Forms A-D are probably due to variations in instructors' pedagogical practices and implementations of the lecture-tutorials.

To push this issue a little farther, note that the majority of classes included in the fall 2010 data were taught by instructors who had not previously participated in this study. Since these classes were scattered across the United States, we could not observe the pedagogical practices of each instructor as we did for the classes taught at the University of Colorado at Boulder. This means we do not know how well these classes' instructors integrated the lecture-tutorials into their daily lessons. We do not know what information they exposed students to in order to prepare them for the lecture-tutorials. We do not know whether the instructors provided enough time for students to work on the lecture-tutorials. We do not know if instructors made students work on the lecture-tutorials in collaborative groups of two to three people. We do not know if instructors simply provided answers to students or asked guiding questions to help the students construct their answers. We do not know whether these instructors used other forms of interactive engagement and, if so, how they were implemented (after all, the non-lecture-tutorial classes may have fostered student learning through other interactive engagement techniques). We do not know if and how the instructors assessed students' masteries of the lecture-tutorials' contents. We do not know whether students in these classes perceived the cosmology lecture-tutorials to be an integral part of the day's lesson or just an activity that was "tacked on." In short, we do not know if the lecture-tutorials were implemented according to the best practices elucidated in Brogt (2007) and Prather *et al.* (2005). Describing and understanding what happens in courses such as these is an important topic for future AER studies, as discussed below.

9.2 Implications for Future Research

What affect might this work have on future AER studies? I foresee four different kinds of research projects that might be influenced by this dissertation.

First, future AER studies may wish to examine other topics in cosmology. While we tried to focus on some of the major areas, there remain many others that we have not addressed. To take just one example, astronomers have a plethora of data on the existence and nature of dark matter. Our study covered just a single piece of this evidence. Future studies may wish to probe students' understandings of other pieces of evidence for dark matter and/or develop interventions and resources, such as (but not limited to) lecture-tutorials, that address any common difficulties. These studies may be conducted at the Astro 101 level, or they may focus on more advanced courses.

Second, this work may provide an important starting point for researchers interested in developing a cosmology concept inventory. Such concept inventories already exist for a variety of astronomical topics, including star properties (Bailey 2007), the greenhouse effect (Keller 2009), lunar phases (Lindell 2001), and light and spectroscopy (Bardar *et al.* 2006). While concept inventories might be limited in the amount of information they can provide (Wallace and Bailey 2010), they are easy to administer and interpret, and have played major roles in shaping our understanding of the strengths and weaknesses of various instructional practices (e.g. Hake 1998a; Prather *et al.* 2009). Researchers interested in developing a cosmology concept inventory may benefit from this study in two ways. First, the common student difficulties uncovered by this study can help shape the distractors of several items on a concept inventory. Second, much of our work in designing, analyzing, and revising our conceptual surveys can save concept inventory developers some time and effort, since we have already found item wordings and topics that are prone to misinterpretation. Much of our work on our surveys mimics the initial stages of the development of a concept inventory.

Third, researchers may wish to study other aspects of students, aside from their conceptual understanding of cosmology, that might be affected by the lecture-tutorials. As others have noted (e.g., Hake 1998b; Redish 2003), studies like this dissertation do not provide information on students' attitudes and beliefs, metacognitive skills, fluency with multiple representations, understanding of the nature and process of scientific inquiry, and ability to address real-world problems. Future studies may wish to investigate many of these aspects.

Finally, the variation we observed in the gain scores of lecture-tutorial and non-lecture-tutorial populations underscores the importance of studies of instructional and implementation practices. As noted above, prior research indicates that the implementation of interactive engagement activities is the most important factor affecting student learning (Hake 1998a; Hake 1998b; Prather *et al.* 2009; Rudolph *et al.* 2010). Other studies show that while many instructors are aware of and interested in research-validated interactive engagement activities, they may exhibit significant variations in how they implement those activities, even to the point where they disregard research-informed best practices (Dancy and Henderson 2010; Henderson and Dancy 2009; Turpen and Finkelstein 2009). Such implementation differences are sometimes due the unique situations and institutional cultures in which instructors work (Henderson and Dancy 2007), although Turpen and Finkelstein (2009) show that different instructors still have different implementation practices even when they are subject to the same situational constraints. Turpen and Finkelstein (2010) demonstrate that these differences do affect classroom norms and students' perceptions. Although we primarily worked from a cognitive perspective in this project, the studies cited above and others (e.g., Finkelstein 2005; Finkelstein and Pollock 2005; Greeno 2006) suggest that situated models of learning are perhaps better suited for studying and understanding instructors' implementation practices.

However, to fully understand the variation in gain scores, we need more than just studies of how instructors implement the lecture-tutorials. We also need to study the instructional practices of classes that do not use the lecture-tutorials. As Figure 9.1 shows, there are some populations of non-lecture-tutorial students that exhibited larger normalized gains than their lecture-tutorial counterparts. What happened in these classes that allowed those students to do so well? Did they use alternative types of interactive engagement activities to promote students' learning of cosmology? We currently lack the data we need to sufficiently answer these question. Future studies, working from a situated perspective on learning, should investigate these classes and elucidate their effective instructional practices.

Bibliography

[1] Ackerman, T. A., Gierl, M. J., and Walker, C. M. *Educational Measurement: Issues & Practice* **22**, 37–51 (2003).

[2] American Education Research Association, American Psychological Association, and National Council on Measurement in Education. *Standards for Educational and Psychological Testing.* American Education Research Association, Washington, DC, (1999).

[3] Anderson, J. R., Reder, L. M., and Simon, H. A. *Educational Researcher* **25**, 5–11 (1996).

[4] Anderson, J. R., Reder, L. M., and Simon, H. A. In *Brookings Papers of Educational Policy: 1998,* Ravitch, D., editor, 227–78. Brookings Institution, Washington, DC (1998).

[5] Andrich, D. *Medical Care* **42**, I7–16 (2004).

[6] Bailey, J. M. *Astronomy Education Review* **6**, 133–9 (2007).

[7] Bailey, J. M., Coble, K., Cochran, G., Sanchez, R., Larrieu, D., Hayes, V. L., Nickerson, V., and McLin, K. Invited talk at the 2011 Winter Meeting of the American Association of Physics Teachers.

[8] Bailey, J. M. and Slater, T. F. *American Journal of Physics* **73**, 677–85 (2005).

[9] Baker, F. B. and Kim, S. *Item Response Theory: Parameter Estimation Techniques.* Marcel Dekker, Inc., New York, NY, 2nd edition, (2004).

[10] Bardar, E. M., Prather, E. E., Brecher, K., and Slater, T. F. *Astronomy Education Review* **5**, 103–13 (2006).

[11] Benezet, L. P. *Journal of the National Education Association* **24**, 241–4 (1935a).

[12] Benezet, L. P. *Journal of the National Education Association* **24**, 301–3 (1935b).

[13] Benezet, L. P. *Journal of the National Education Association* **25**, 7–8 (1936).

[14] Bennett, J., Donahue, M., Schneider, N., and Voit, M. *The Cosmic Perspective.* Pearson Education, Inc., San Francisco, CA, (2008).

[15] Black, P. *Educational Designer* **1** (2009).
URL: www.educationaldesigner.org/ed/volume1/issue3/article12.

[16] Bond, T. G. and Fox, C. M. *Applying the Rasch Model: Fundamental Measurement in the Human Sciences.* Lawrence Erlbaum Associates, Mahwah, NJ, (2001).

[17] Briggs, D. C., Alonzo, A. C., Schwab, C., and Wilson, M. *Educational Assessment* **11**, 33–63 (2006).

[18] Briggs, D. C. and Wilson, M. *Journal of Applied Measurement* **4**, 87–100 (2003).

[19] Brogt, E. *Astronomy Education Review* **6**, 50–8 (2007).

[20] Brown, D. E. and Clement, J. *Instructional Science* **18**, 237–61 (1989).

[21] Brown, J. S., Collins, A., and Duguid, P. *Educational Researcher* **18**, 32–42 (1989).

[22] Carey, S. In *Ontogeny, Phylogeny, and Historical Development,* Strauss, S., editor, 1–27. Ablex Publishing Corporation, Norwood, NJ (1988).

[23] Cheng, T. *Relativity, Gravitation and Cosmology.* Oxford University Press, Oxford, UK, 2nd edition, (2010).

[24] Clement, J. *American Journal of Physics* **50**, 66–71 (1982).

[25] Clement, J., Brown, D. E., and Zietsman, A. *International Journal of Science Education* **11**, 554–65 (1989).

[26] Coble, K., McLin, K. M., Metevier, A. J., Bailey, J. M., and Cominsky, L. R. Invited talk at the 2011 Winter Meeting of the American Association of Physics Teachers.

[27] Cohen, J. *Educational and Psychological Measurement* **20**, 37–46 (1960).

[28] Cohen, J. *Psychological Bulletin* **70**, 213–20 (1968).

[29] Comins, N. F. *Heavenly Errors: Misconceptions About the Real Nature of the Universe.* Columbia University Press, New York, NY, (2001).

[30] Dancy, M. and Henderson, C. *American Journal of Physics* **78**, 1056–63 (2010).

[31] Davis, T. M. and Lineweaver, C. H. *Publications of the Astronomical Society of Australia* **21**, 97–109 (2004).

[32] Deming, G. and Hufnagel, B. *The Physics Teacher* **39**, 368–9 (2001).

[33] Ding, L. and Beichner, R. *Physical Review Special Topics – Physics Education Research* **5**, 020103 (2009).

[34] Ding, L., Chabay, R., Sherwood, B., and Beichner, R. *Physical Review Special Topics – Physics Education Research* **2**, 010105 (2006).

[35] diSessa, A. A. *Cognition & Instruction* **2 & 3**, 105–225 (1993).

[36] diSessa, A. A. and Sherin, B. L. *International Journal of Science Education* **20**, 1155–91 (1998).

[37] Dodelson, S. *Modern Cosmology.* Academic Press, San Diego, CA, (2003).

[38] Duncan, D. *Astronomy Education Review* **5**, 70–88 (2007).

[39] Duncan, T. and Tyler, C. *Your Cosmic Context: An Introduction to Modern Cosmology.* Pearson Education, Inc., San Francisco, CA, (2009).

[40] Elby, A. *The Journal of Mathematical Behavior* **19**, 481–502 (2000).

[41] Elby, A. *American Journal of Physics* **69**, S54–64 (2001).

[42] Embretson, S. E. and Reise, S. P. *Item Response Theory for Psychologists.* Lawrence Erlbaum Association, Inc., Mahwah, NJ, (2000).

[43] Erickson, F. In *Handbook of Research on Teaching, 3rd Edition*, 119–161. Macmillian Publishing Company, New York, NY (1986).

[44] Finkelstein, N. *Internation Journal of Science Education* **27**, 1187–1209 (2005).

[45] Finkelstein, N. D. and Pollock, S. J. *Physical Review Special Topics – Physics Education Research* **1**, 010101 (2005).

[46] Fixen, D. J., Cheng, E. S., Gales, J. M., Mather, J. C., Shafer, R. A., and Wright, E. L. *The Astrophysicsl Journal* **473**, 576–87 (1996).

[47] Fleiss, J. L., Levin, B., and Paik, M. C. *Statistical Methods for Rates and Proportions.* John Wiley and Sons, Inc., Hoboken, NJ, 3rd edition, (2003).

[48] Fraknoi, A. *Astronomy Education Review* **1**, 121–3 (2002).

[49] Frank, B. W., Kanim, S. E., and Gomez, L. S. *Physical Review Special Topics – Physics Education Research* **4**, 020102 (2008).

[50] George, D. and Mallory, P. *SPSS for Windows Step by Step: A Simple Guide and Reference.* Pearson Education, Boston, MA, (2009).

[51] Gorin, J. S. *Educational Measurement: Issues and Practice* **25**, 21–35 (2006).

[52] Green, P. J. *Peer Instruction for Astronomy.* Pearson Education, Inc., Upper Saddle River, NJ, (2003).

[53] Greeno, J. G. *Educational Researcher* **26**, 5–17 (1997).

[54] Greeno, J. G. In *Learning in Activity,* Sawyer, R. K., editor, 79–96. Cambridge University Press, New York, NY (2006).

[55] Guilford, J. P. and Fruchter, B. *Fundamental Statistics in Psychology and Education.* McGraw-Hill, Inc., New York, NY, 6th edition, (1978).

[56] Hake, R. R. *American Journal of Physics* **66**, 64–74 (1998a).

[57] Hake, R. R. *unpublished* (1998b).

[58] Halloun, I. and Hestenes, D. *American Journal of Physics* **53**, 1043–55 (1985).

[59] Hambleton, R. K. and Jones, R. J. *Educational Measurement: Issues & Practice* **12**, 253–62 (1993).

[60] Hammer, D., Elby, A., Scherr, R. E., and Redish, E. F. In *Transfer of Learning from a Modern Multidisciplinary Perspective,* Mestre, J., editor, 89–120. Information Age Publishing, Scottsdale, AZ (2005).

[61] Harris, D. *Educational Measurement: Issues & Practice* **8**, 157–63 (1989).

[62] Hawking, S. *A Brief History of Time.* Bantam Books, New York, (1988).

[63] Heller, P. and Hollabaugh, M. *American Journal of Physics* **60**, 637–44 (1992).

[64] Heller, P., Keith, R., and Anderson, S. *American Journal of Physics* **60**, 627–36 (1992).

[65] Henderson, C. and Dancy, M. H. *Physical Review Special Topics – Physics Education Research* **3**, 020102 (2007).

[66] Henderson, C. and Dancy, M. H. *Physical Review Special Topics – Physics Education Research* **5**, 020107 (2009).

[67] Henriques, L., Colburn, A., and Ritz, W. C. In *Developing Assessment Items: A How-To Guide,* McMahon, M., Simmons, P., Sommers, R., DeBaets, D., and Crawley, F., editors, 15–30. NSTA Press, Arlington, VA (2006).

[68] Heron, P. R. L. In *Proceedings of the International School of Physics "Enrico Fermi" Course CLVI,* Redish, E. F. and Vicentini, M., editors, 341–50. IOP Press, Amsterdam, Netherlands (2004a).

[69] Heron, P. R. L. In *Proceedings of the International School of Physics "Enrico Fermi" Course CLVI,* Redish, E. F. and Vicentini, M., editors, 351–65. IOP Press, Amsterdam, Netherlands (2004b).

[70] Hestenes, D. and Wells, M. *The Physics Teacher* **30**, 159–66 (1992).

[71] Hestenes, D., Wells, M., and Swackhamer, G. *The Physics Teacher* **30**, 141–58 (1992).

[72] Hogg, D. W. . astro-ph/9905116v4 (2000).

[73] Hubble, E. *Proceedings of the National Academies of Sciences* **168-73**, 1–22 (1929).

[74] Hudgins, D. W., Prather, E. E., Grayson, D. J., and Smits, D. P. *Astronomy Education Review* **5**, 1–22 (2006).

[75] Hufnagel, B. *Astronomy Education Review* **1**, 47–51 (2001).

[76] Kane, M. T. *Psychological Bulletin* **112**, 527–35 (1992).

[77] Keller, J. M. *Development of a Concept Inventory Addressing Students' Beliefs and Reasoning Difficulties Regarding the Greenhouse Effect.* PhD thesis, University of Arizona, Tucson, AZ, USA, (2009).

[78] Kohl, P. B. and Finkelstein, N. D. *Physical Review Special Topics – Physics Education Research* **2**, 010102 (2006).

[79] Kohl, P. B., Rosengrant, D., and Finkelstein, N. D. *Physical Review Special Topics – Physics Education Research* **3**, 010108 (2007).

[80] Kohl, P. B. and Finkelstein, N. D. *Physical Review Special Topics – Physics Education Research* **4**, 010111 (2008).

[81] Kolb, E. W. and Turner, M. S. *The Early Universe.* Addison-Wesley Publishing Company, Readwood City, CA, (1990).

[82] Landis, J. R. and Koch, G. G. *Biometrics* **33**, 159–74 (1977).

[83] Ledoux, C., Petitjean, P., and Srianand, R. *The Astrophysical Journal* **640**, L25–8 (2006).

[84] Lewis, A. D., Boute, D. A., and Stocke, J. T. *The Astrophysical Journal* **586**, 135–42 (2003).

[85] Lightman, A. P., Miller, J. D., and Leadbeater, B. J. In *Proceedings of the Second International Seminar Misconceptions and Educational Strategies in Science and Mathematics, Volume III,* Novak, J. D., editor, 616–41. Cornell University Press, Ithaca, NY (1987).

[86] Lightman, A. P. and Miller, J. D. *Social Studies of Science* **19**, 127–36 (1989).

[87] Lightman, A. and Sadler, P. *The Physics Teacher* **31**, 162–7 (1993).

[88] Lindell, R. S. *Enhancing College Students' Understanding of Lunar Phases.* PhD thesis, University of Nebraska, Lincoln, NE, USA, (2001).

[89] Lineweaver, C. H. and Davis, T. M. *Scientific American* **292**, 36–45 (2005).

[90] LoPresto, M. C. and Murrell, S. R. *Astronomy Education Review* **8**, 010105 (2009).

[91] Lord, F. M. and Novick, M. R. *Statistical Theories of Mental Test Scores.* Addison-Wesley Publishing Company, Reading, MA, (1968).

[92] Lyman, F. T. In *Mainstreaming Digest,* Anderson, A., editor, 109–13. University of Maryland Press, College Park, MD (1981).

[93] Maloney, D. P., O'Kuma, T. L., Heiggelke, C. J., and van Heuvelen, A. *American Journal of Physics* **69**, S12–23 (2001).

[94] Masters, G. N. *Psychometrika* **47**, 149–74 (1982).

[95] Masters, G. N. *Journal of Educational Measurement* **25**, 15–29 (1988).

[96] Mazur, E. *Peer Instruction: A User's Manual.* Prentice Hall, Inc., Upper Saddle River, NJ, (1997).

[97] McCloskey, M., Caramazza, A., and Green, B. *Science* **210**, 1139–41 (1980).

[98] McCloskey, M., Washburn, A., and Felch, L. *Journal of Experimental Psychology: Learning, Memory, and Cognition* **9**, 636–49 (1983).

[99] McClymer, J. F. and Knoles, L. Z. *Journal on Excellence in College Teaching* **3**, 33–50 (1992).

[100] McDermott, L. C. *American Journal of Physics* **59**, 301–15 (1991).

[101] McDermott, L. C. and Redish, E. F. *American Journal of Physics* **67**, 755–67 (1999).

[102] McDermott, L. C., Rosenquist, M. L., and van Zee, E. H. *American Journal of Physics* **55**, 503–13 (1987).

[103] McDermott, L. C. and Shaffer, P. S. *American Journal of Physics* **60**, 994–1003 (1992a).

[104] McDermott, L. C. and Shaffer, P. S. *American Journal of Physics* **60**, 1003–13 (1992b).

[105] McDermott, L. C., Shaffer, P. S., and the Physics Education Group. *Tutorials in Introductory Physics.* Pearson Education, Inc., Upper Saddle River, (2002).

[106] Meltzer, D. E. and Manivannan, K. *American Journal of Physics* **70**, 639–54 (2002).

[107] Minstrell, J. In *Research in Physics Learning: Theoretical Issues and Empirical Studies: Proceedings of an International Workshop Held at the University of Bremen, March 4-8, 1991,* Duit, R., Goldberg, F., and Niedderer, H., editors, 110–28. Institut für die Pädagogik der Naturwissenschaften an der Universität Kiel, Kiel, Germany (1992).

[108] National Research Council. *How People Learn: Brain, Mind, Experience, and School.* National Academy of Sciences, Washington, DC, expanded edition, (2000).

[109] Otero, V., Finkelstein, N., McCray, R., and Pollock, S. *Science* **313**, 445–6 (2006).

[110] Otero, V. K. and Harlow, D. B. In *Reviews in PER Vol. 2: Getting Started in PER,* Henderson, C. and Harper, K. A., editors, 1–66. American Association of Physics Teachers, College Park, MD (2009).

[111] Otero, V., Pollock, S., and Finkelstein, N. *American Journal of Physics* **78**, 1218–24 (2010).

[112] Patton, M. Q. *Qualitative Evaluation Methods.* SAGE Publications, Thousand Oaks, CA, (1980).

[113] Peebles, P. J. E. *Principles of Physical Cosmology.* Princeton University Press, Princeton, NJ, (1993).

[114] Perlmutter, S., Aldering, G., Goldhaber, G., Knop, R. A., Nugent, P., Castro, P. G., Deustua, S., Fabbro, S., Goobar, A., Groom, D. E., Hook, I. M., Kim, A. G., Kim, M. Y., Lee, J. C., Nunes, N. J., Pain, R., Pennypacker, C. R., and Quimby, R. *The Astrophysical Journal* **517**, 565–86 (1999).

[115] Posner, G. J., Strike, K. A., Hewson, P. W., and Gertzog, W. A. *Science Education* **66**, 211–27 (1982).

[116] Prather, E. E. and Brissenden, G. *Astronomy Education Review* **7**, 1–17 (2009).

[117] Prather, E. E., Rudolph, A. L., Brissenden, G., and Schlingman, W. M. *American Journal of Physics* **77**, 320–30 (2009).

[118] Prather, E. E., Slater, T. F., Adams, J. P., Bailey, J. M., Jones, L. V., and Dostal, J. A. *Astronomy Education Review* **3**, 122–36 (2005).

[119] Prather, E. E., Slater, T. F., Adams, J. P., Brissenden, G., Dostal, J. A., and Team, C. *Lecture-Tutorials for Introductory Astronomy.* Pearson Education, Inc., San Francisco, CA, 2nd edition, (2008).

[120] Prather, E. E., Slater, T. F., and Offerdahl, E. G. *Astronomy Education Review* **1**, 28–34 (2003).

[121] Primack, J. R. *New Astronomy Reviews* **49**, 25–34 (2005).

[122] Rasch, G. *Probabilistic Models for Some Intelligence and Attainment Tests.* University of Chicago Press, Chicago, IL, (1960).

[123] Redish, E. F. *American Journal of Physics* **62**, 796–803 (1994).

[124] Redish, E. F. *Teaching Physics with the Physics Suite.* John Wiley & Sons, Inc., Hoboken, NJ, (2003).

[125] Reiss, A. G., Filippenko, A. V., Challis, P., Clocchiatti, A., Diercks, A., Garnavich, P. M., Gilliland, R. L., Hogan, C. J., Jha, S., Kirshner, R. P., Leibundgut, B., Phillips, M. M., Reiss, D., Schmidt, B. P., Schommer, R. A., Smith, R. C., Spyromilio, J., Stubbs, C., Suntzeff, N. B., and Tonry, J. *The Astrophysical Journal* **116**, 1009–38 (1998).

[126] Rubin, V. C., Ford, W. K., and Thonnard, N. *The Astrophysical Journal* **238**, 471–87 (1980).

[127] Rudolph, A. L., Prather, E. E., Brissenden, G., Consiglio, D., and Gonzaga, V. *Astronomy Education Review* **9**, 010107 (2010).

[128] Sanders, R. H. and McGaugh, S. S. *Annual Reviews of Astronomy and Astrophysics* **40**, 263–317 (2002).

[129] Scherr, R. E., Shaffer, P. S., and Vokos, S. *American Journal of Physics* **69**, S24–35 (2001).

[130] Scherr, R. E., Shaffer, P. S., and Vokos, S. *American Journal of Physics* **70**, 1238–48 (2002).

[131] Schmitt, N. *Psychological Assessment* **8**, 350–3 (1996).

[132] Schoenfeld, A. H. In *Cognitive Science and Mathematics Education,* Schoenfeld, A. H., editor, 189–215. Lawrence Erlbaum Associates, Mahwah, NJ (1987).

[133] Schoenfeld, A. H. *Educational Designer* **1** (2009).
URL: www.educationaldesigner.org/ed/volume1/issue2/article5.

[134] Schunn, C. *Educational Designer* **1** (2008).
URL: www.educationaldesigner.org/ed/volume1/issue1/article2.

[135] Sfard, A. *Educational Researcher* **27**, 4–13 (1998).

[136] Shepard, L. A. In *Review of Research in Education,* Darling-Hammond, L., editor, 405–50. American Educational Research Association, Washington, DC (1993).

[137] Shepard, L. A. *Educational Researcher* **29**, 4–14 (2000).

[138] Siegel, M. A., Wissehr, C., and Halverson, K. *The Science Teacher* **75**, 43–6 (2008).

[139] Simonelli, G. and Pilachowski, C. A. *Astronomy Education Review* **2**, 166–71 (2004).

[140] Slater, T., Adams, J. P., Brissenden, G., and Duncan, D. *The Physics Teacher* **39**, 52–5 (2001).

[141] Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., Guild, N., and Su, T. T. *Science* **323**, 122–4 (2009).

[142] Smith, R. M. and Miao, C. Y. In *Objective Measurement: Theory Into Practice,* Wilson, M., editor, volume 2, 316–27. Ablex Publishing Corporation, Norwood, NJ (1994).

[143] Sofue, Y. and Rubin, V. *Annual Reviews of Astronomy and Astrophysics* **39**, 137–74 (2001).

[144] Sokoloff, D. R. and Thornton, R. K. *The Physics Teacher* **35**, 340–7 (1997).

[145] Sparke, L. S. and Gallagher, J. S. *Galaxies in the Universe: An Introduction.* Cambridge University Press, New York, NY, (2000).

[146] Spergel, D. N., Bean, R., Doré, O., Nolta, M. R., Bennett, C. L., Dunkley, J., Hinshaw, G., Jarosik, N., Komatsu, E., Page, L., Peiris, H. V., Verde, L., Halpern, M., Hill, R. S., Kogut, A., Limon, M., Meyer, S. S., Odegard, N., Tucker, G. S., Weiland, J. L., Wollack, E., and Wright, E. L. *The Astrophysical Journal Supplement Series* **170**, 377–408 (2007).

[147] Stevens, S. S. *Science* **103**, 677–80 (1946).

[148] Strike, K. A. and Posner, G. J. In *Cognitive Structure and Conceptual Change,* West, L. H. T. and Pines, A. L., editors, 211–31. Academic Press, Inc., Orlando, FL (1985).

[149] Strike, K. A. and Posner, G. J. In *Philosophy of Science, Cognitive Psychology, and Educational Theory and Practice,* Duschl, R. A. and Hamilton, R. J., editors, 147–76. State University of New York Press, Albany, NY (1992).

[150] Swan, M. *Educational Designer* **1** (2008).
URL: www.educationaldesigner.org/ed/volume1/issue1/article3.

[151] Taylor, J. R. *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements.* University Science Books, Mill Valley, CA, (1982).

[152] Tegmark, M., Zaldarriaga, M., and Hamilton, A. J. S. *Physical Review D* **63**, 043007 (2001).

[153] Thompson, B. In *Score Reliability,* Thompson, B., editor, 3–30. SAGE, Thousand Oaks, CA (2003).

[154] Trowbridge, D. E. and McDermott, L. C. *American Journal of Physics* **48**, 1020–8 (1980).

[155] Trowbridge, D. E. and McDermott, L. C. *American Journal of Physics* **49**, 242–53 (1981).

[156] Turpen, C. and Finkelstein, N. D. *Physical Review Special Topics – Physics Education Research* **5**, 020101 (2009).

[157] Turpen, C. and Finkelstein, N. D. *Physical Review Special Topics – Physics Education Research* **6**, 020123 (2010).

[158] Van Heuvelen, A. *American Journal of Physics* **59**, 891–7 (1991).

[159] Vosniadou, S. *Learning & Instruction* **4**, 45–69 (1994).

[160] Vygotsky, L. S. In *Readings on the Development of Children,* Gauvain, M. and Cole, M., editors, 34–40. Scientific American Books, New York, NY (1978).

[161] Wainer, H. and Thissen, D. In *Test Scoring,* Wainer, H. and Thissen, D., editors, 23–72. Lawrence Erlbaum Associates, Mahwah, NJ (2001).

[162] Wallace, C. S. and Bailey, J. M. *Astronomy Education Review* **9**, 010116 (2010).

[163] Wiggins, G. *Educative Assessment.* Jossey-Bass Publishers, San Francisco, CA, (1998).

[164] Wilcox, R. R. *New Statistical Procedures for the Social Sciences: Modern Solutions to Basic Problems.* Lawrence Erlbaum Associates, Hillsdale, NJ, (1987).

[165] Willis, G. D. *Cognitive Interviewing: A Tool for Improving Questionnaire Design.* SAGE Publications, Thousand Oaks, CA, (2005).

[166] Wilson, M. *Constructing Measures: An Item Response Modeling Approach.* Lawrence Erlbaum Associates, Inc., Mahwah, NJ, (2005).

[167] Wright, B. D. *Educational Measurement: Issues & Practice* **16**, 33–45 (1997).

[168] Wright, B. D. and Linacre, J. M. *Archives of Physical Medicine and Rehabilitation* **70**, 857–90 (1989).

[169] Wu, M. and Adams, R. J. *Journal of Applied Measurement* , in press (2011).

[170] Yen, W. M. *Applied Psychological Measurement* **8**, 125–45 (1984).

[171] Yen, W. M. *Journal of Educational Measurement* **30**, 187–213 (1993).

[172] Yen, W. M. and Fitzpatrick, A. R. In *Educational Measurement, 4th edition,* Brennan, R., editor, 111–53. American Council on Education/Praeger, Westport, CT (2006).

[173] Zeilik, M., Schau, C., Mattern, N., Hall, S., Teague, K. W., and Bisard, W. *American Journal of Physics* **65**, 987–96 (1997).

[174] Zwicky, F. *The Astrophysical Journal* **86**, 217–46 (1937).

Appendix A

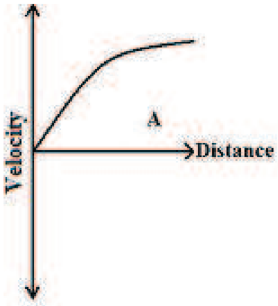
Fall 2009 Surveys

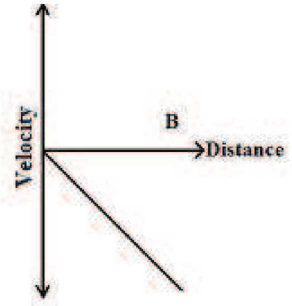
A.1 Form A

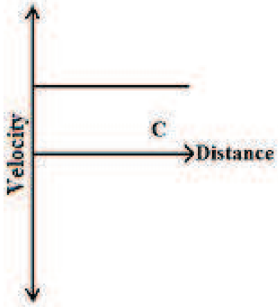
The fall 2009 version of Form A begins on the next page.

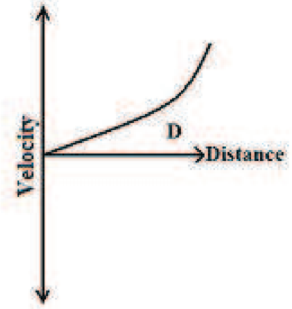
Name: ______________________________ **Conceptual Cosmology Survey - Form A**

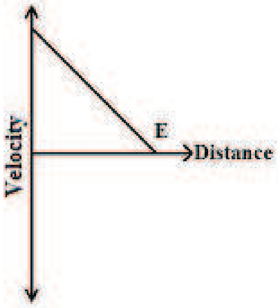
1) Below are some possible plots (A - H) of how fast galaxies are moving versus their distance away from us.

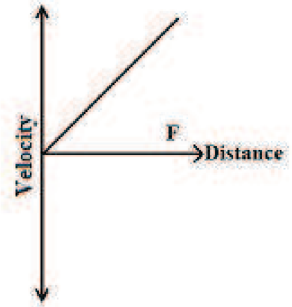


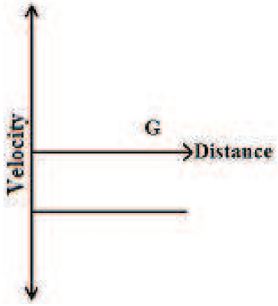


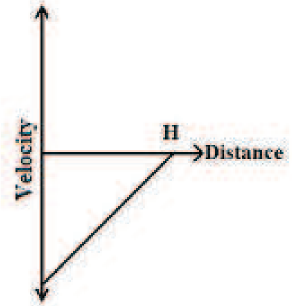












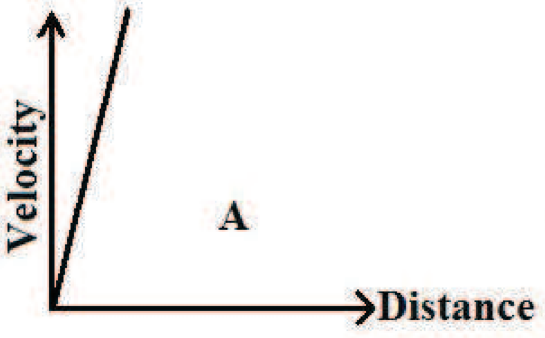
Which graph or graphs show a universe that is expanding at a constant rate? ***Explain your reasons for making your selection(s).***

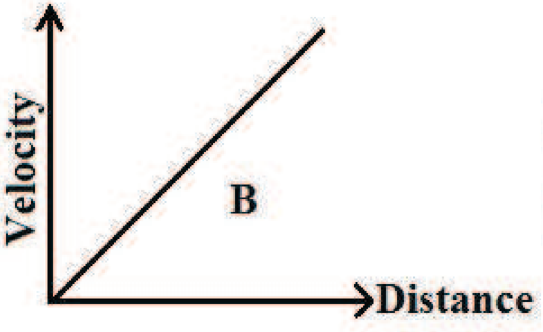
Which graph or graphs show a universe that is contracting at a constant rate? ***Explain your reasons for making your selection(s).***

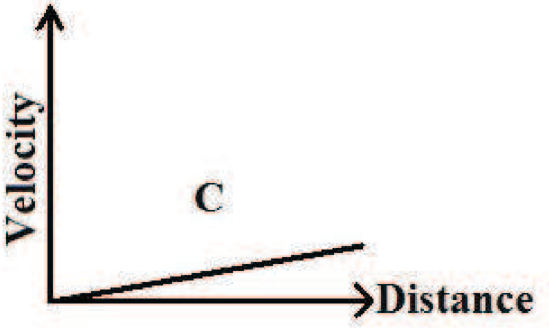
Which graph or graphs show a universe that is expanding at a faster and faster rate over time? ***Explain your reasons for making your selection(s).***

1) (con't) Which graph or graphs on the previous page show a universe that is expanding at a slower and slower rate over time? ***Explain your reasons for making your selection(s).***

2) Below are some plots (A - C) of how fast galaxies are moving versus their distance from us.







Which of these three plots predicts an **old** universe? ***Explain your reasons for making your selection.***

A.2 Form B

The fall 2009 version of Form B begins on the next page.

Name: ___________________________________ **Conceptual Cosmology Survey - Form B**

1) Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." ***Provide a drawing if possible to help illustrate your thinking.***

2) Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory." ***Provide a drawing if possible to help illustrate your thinking.***

3) Imagine you had a spaceship that could take you back in time to the time of the Big Bang. Would there have been any locations in the universe from which you could have watched the Big Bang from a distance? *If yes*, describe what it would have looked like and draw a picture. *If no*, explain why not.

4) If you could travel to any location in the universe, could you ever see the center of the universe? ***Explain your reasoning and provide a drawing if possible to help illustrate your thinking.***

5) If you could travel to any location in the universe, could you go to a place where there would be no galaxies in front of you? ***Explain your reasoning and provide a drawing if possible to help illustrate your thinking.***

6) Which of the following is **true**? ***There may be more than one correct answer. Select all that apply.***

a) Over time, the distances between widely separated planets in the solar system **always increase** and **never decrease**.
b) Over time, the distances between widely separated stars in the galaxy **always increase** and **never decrease**.
c) Over time, the distances between widely separated galaxies in the universe **always increase** and **never decrease**.

Explain your reasoning for your choice(s).

A.3 Form C

The fall 2009 version of Form C begins on the next page.

Name: ____________________________________ **Conceptual Cosmology Survey - Form C**

1) Galaxy X and Galaxy Y are currently 8 billion light-years apart in the expanding universe. How long will light from Galaxy X take to reach Galaxy Y?

   a) less than 8 billion years b) exactly 8 billion years
   c) more than 8 billion years d) there is not enough information to tell

   Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

2) Imagine you lived in the universe 10 billion years ago.
   Was the star formation rate of the universe higher, lower, or about the same as it is today? ***Explain your reasoning.***

   Was the fraction of galaxies that were elliptical higher, lower, or about the same as it is today? ***Explain your reasoning.***

   As time went on, did the fraction of active galaxies go up, go down, or did it stay about the same? ***Explain your reasoning.***

3) Has the temperature of the universe changed over time, or has it always been about the same? ***Explain your reasoning and provide a drawing if possible to help illustrate your thinking.***

4) How does the total amount of matter in the universe ***right now*** compare to the total amount of matter in the universe at the ***very beginning*** of the universe (the moment just after the Big Bang)? ***Explain your reasoning and provide a drawing if possible to help illustrate your thinking.***

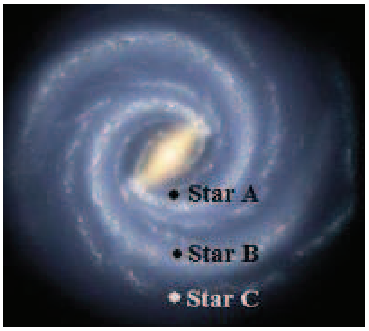
5) Has the density of matter in the universe changed over time, or has it always been about the same? ***Explain your reasoning and provide a drawing if possible to help illustrate your thinking.***

A.4 Form D

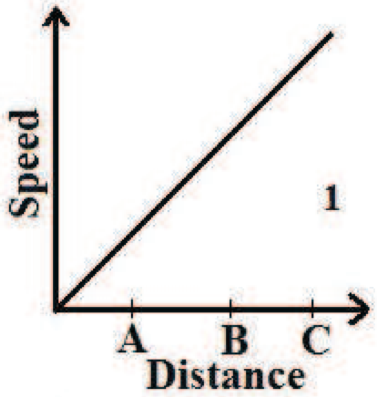
The fall 2009 version of Form D begins on the next page.

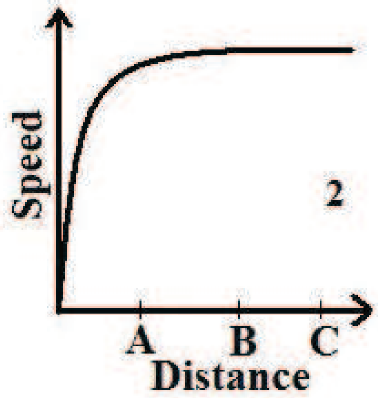
Name: ______________________________________ **Conceptual Cosmology Survey - Form D**

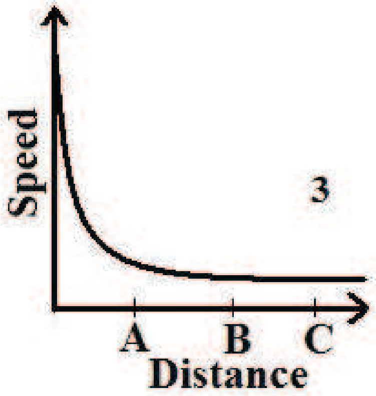
1) Here is a picture of a spiral galaxy. The locations of three stars (A - C) are shown.

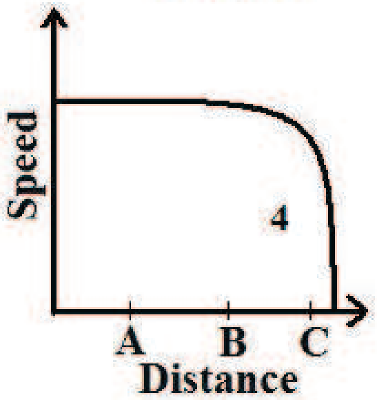


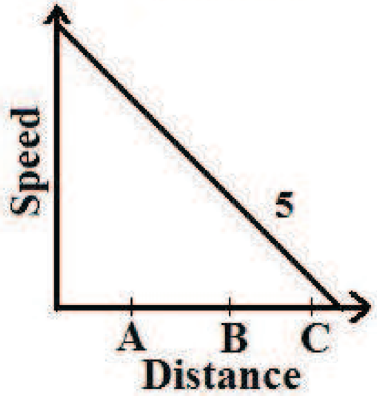
When we try to observe how fast stars orbit the center of the galaxy (especially far from the galaxy's center), we draw what's called a ***rotation curve*** for the galaxy. Below are some possible rotation curves (1-6). The locations of stars A-C are shown on the distance axis.

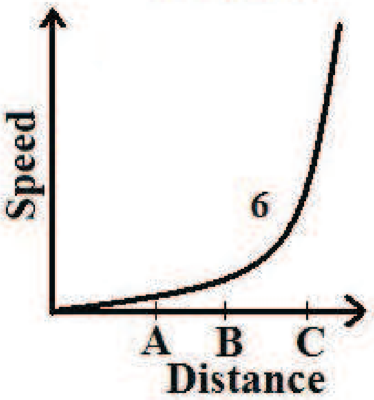












Which one graph best represents how objects actually move in our galaxy? ***Explain the reasoning behind your choice.***

2) Based on your selection in Question 1, describe how the speeds at which stars A, B, and C orbit the galaxy compare to one another.

3) Based on your previous answers, how would the net gravitational force exerted on stars A, B, and C compare? ***Explain your reasoning.***

4) Based on your previous answers, what does this graph tell us about how matter is distributed in our galaxy? ***Explain your reasoning.***

Appendix B

Fall 2009 Lecture-Tutorials

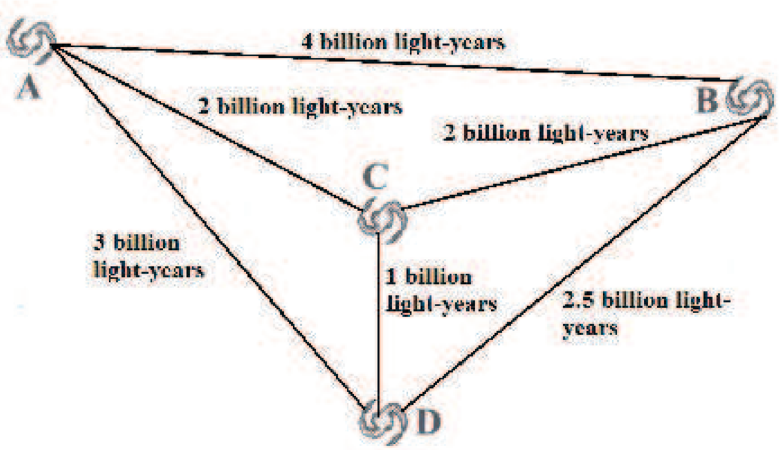
B.1 Hubble's Law

The fall 2009 version of the "Hubble's Law" lecture-tutorial begins on the next page.

Hubble's Law

Consider the small section of the universe containing four galaxies (A-D), shown in Figure 1 below. The distances between each galaxy are also shown.

Figure 1



1) Imagine that this section of the universe doubles in size over time due to the expansion of the universe. Draw what the above section of the universe would look like after it doubles in size. Be sure to identify the new distances between the galaxies.

Hubble's Law

2) Which of the galaxies (B-D) increased its distance from Galaxy A by the greatest amount during this time? Explain your reasoning.

3) Two students are discussing their answers to Question 2:

Student 1: *All of the distances doubled, so all of the distances increased by the same amount. There is no one galaxy who's distance from Galaxy A increased the most.*

Student 2: *Your right that all the distances double in size, but I don't agree that they all increase by the same number of light-years. Since Galaxy D was the farthest away from Galaxy A initially, its distance will increase by the greatest amount when this section of the universe doubles in size.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

4) Describe the relationship between a galaxy's distance from Galaxy A and the speed at which it appears to be moving away from Galaxy A.

5) Is the relationship you described in Question 4 unique to Galaxy A, or would you observe the same relationship if you lived in one of the other galaxies? Explain your reasoning.

The relationship you described in Questions 4 and 5 is called *Hubble's law*. We can depict Hubble's law with the graph shown below. This graph plots the speed at which a galaxy appears to move away from us versus its distance from us. This type of graph is called a *Hubble plot*. Each dot on the plot represents a different galaxy.

Hubble's Law

Figure 2

6) Explain how the Hubble plot shown in Figure 2 above is consistent with the relationship you described in Question 4.

7) Imagine the Hubble plot shown in Figure 2 represents a universe that is continuously doubling in size. Which of the two Hubble plots shown in Figures 3 and 4 below might represent a universe that is tripling in size over the same amount of time? Explain your reasoning.

Figure 3

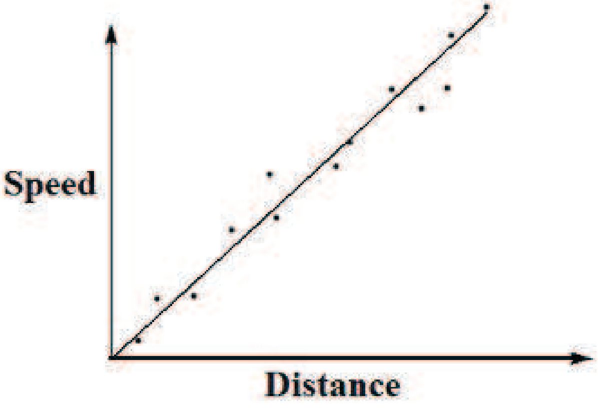
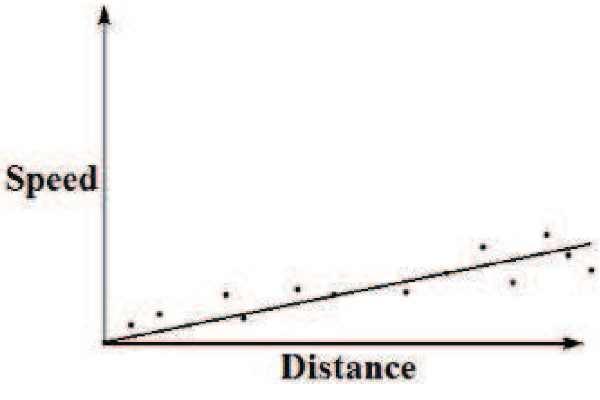


Figure 4



Hubble's Law

The *expansion rate* of the universe determines how fast it is growing in size. For example, a universe that is tripling in size is will have a faster *expansion rate* than a universe that is doubling in size over the same amount of time. In a Hubble Plot, the *expansion rate* is indicated by the slope of the graph.

9) Would you say the expansion rate for the universe represented in Figure 2 is constant, increasing, or decreasing?

10) Rank the expansion rates (from fastest to slowest) of the three different universes represented in Figures 2, 3, and 4. Explain your reasoning.

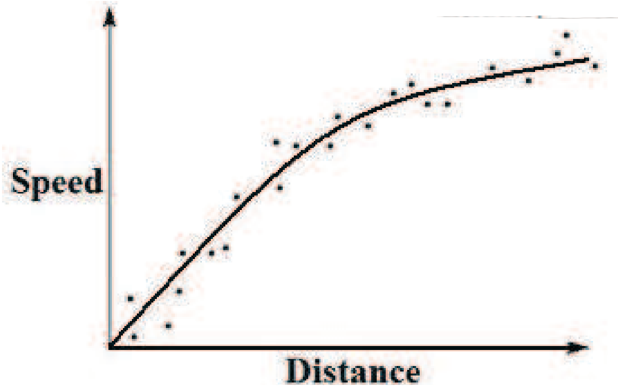
11) If the expansion rate of our universe had been faster, would the universe have reached its current size earlier in its history or later? Explain your reasoning.

12) Complete the blanks in the following sentence using the words provided in parentheses ().

A universe with a fast expansion rate will take a _________ (longer/shorter) amount of time to reach its current size as compared to a universe with a slow expansion rate. If two universes are the same size, but one has a faster expansion rate than the other, then the one with the fast expansion rate must be __________(younger/older) than the one with the slow expansion rate.

Recent observations indicate the Hubble plot for our universe looks like the plot in Figure 5.

Figure 5



Hubble's Law

13) Based on the Hubble Plot in Figure 5, is the expansion rate represented by the motion of galaxies far away from us faster than, slower than, or the same as the expansion rate represented by motion of nearby galaxies? Explain your reasoning.

14) Based on the Hubble Plot in Figure 5, and your answer to the previous question, would you say that the expansion rate of our universe is constant, increasing, or decreasing over time? Explain your reasoning.

15) Consider the following debate between two students regarding their answer to the previous question:

Student 1: *The expansion rate of our universe must be slowing down as time goes on. If you look at the Hubble plot, you can see that the farther away you look, the slower the expansion rate is. The rate at which the most distant galaxies are moving away from us has started to slow down and eventually the expansion rate of nearby galaxies will also slow down.*

Student 2: *I think you are reading the graph wrong. The graph shows that the expansion rate is slower for distant galaxies than it is for nearby galaxies. But the farther we look into space, the further we are looking back in time. Since the Hubble plot is flatter in the past and steeper for nearby galaxies that means the expansion rate has since sped over time*

Do you agree or disagree with either or both of the students? Explain your reasoning.

Hubble's Law

More distant galaxies appear dimmer in our sky than closer galaxies. This is because the amount of light we receive from a galaxy decreases as its distance increases. So the farther away a galaxy is from us, the harder it is to see.

16) If this trend shown in Figure 5 continues, will we have an increasingly easier or harder time seeing other galaxies as the universe ages? Explain your reasoning.

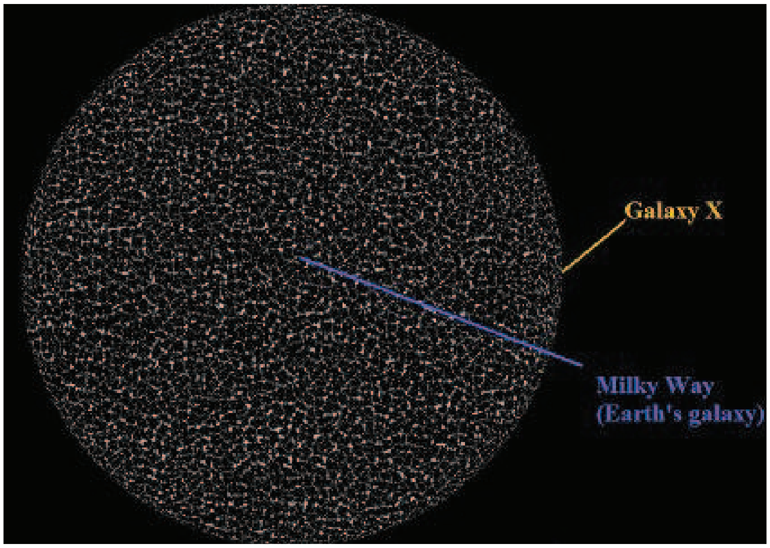
B.2 Making Sense of the Universe and Expansion

The fall 2009 version of the “Making Sense of the Universe and Expansion” lecture-tutorial begins on the next page.

Making Sense of the Universe and Expansion

Part I: The Observable Universe

Each dot in the picture below represents a galaxy. The Milky Way galaxy is represented by the dot at the center of the picture. All of the galaxies inside the circle can be seen from Earth. The circumference of this circle defines what is called our *observable universe*. Any galaxy that exists outside the circle is so far away that its light has not had time to reach Earth and is therefore not part of our observable universe.



1) Do you think the galaxies we can see from Earth are the only galaxies in the universe? Explain your reasoning.

2) Draw a circle around Galaxy X that represents its observable universe.

3) Is the observable universe that you drew for Galaxy X different in size than the observable universe for the Earth? Explain your reasoning.

Making Sense of the Universe and Expansion

4) Two students are talking about the observable universe for Galaxy X:

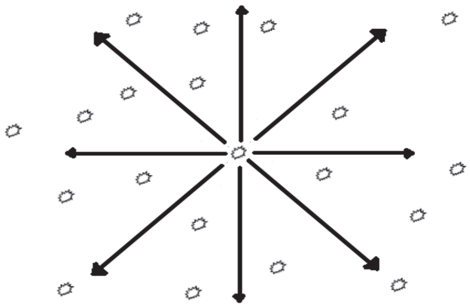
Student 1: *People living in Galaxy X have a strange view of the universe. When they look in one direction they see a bunch of galaxies, but when they look in the other direction all they see is empty space. Galaxy X must be at the edge of the universe since there's nothing but black, empty space beyond it. We're lucky we live at the center since we can see galaxies all the way out to the edge of the universe, no matter where we look.*

Student 2: *I think you're wrong. People living in Galaxy X would probably see a bunch of galaxies in every direction they look, but they can see some galaxies that we can't. The observable universe for any galaxy should look the same as ours. I don't think we are at the center of the universe and I don't think Galaxy X is at the edge either*

Do you agree or disagree with either or both of the students? Explain your reasoning.

Part II: An Analogy for Expansion

Consider the following description and sketch from a student regarding the expansion of the universe.

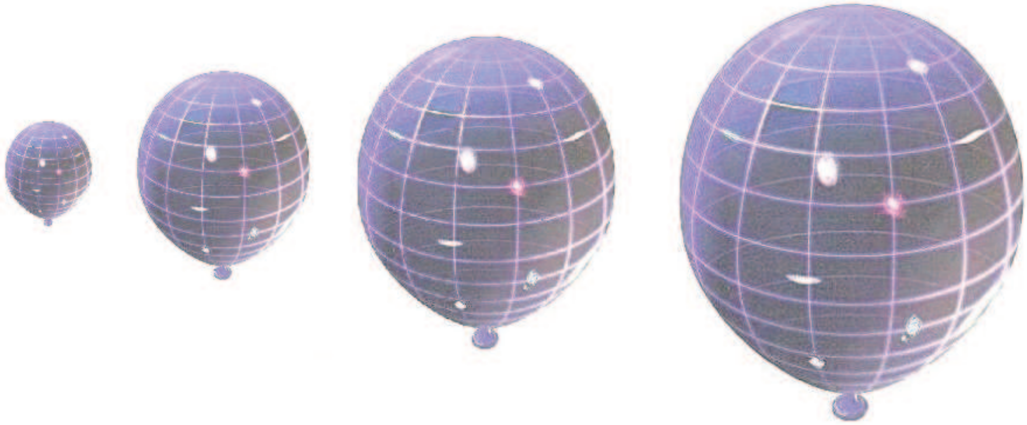


Student: *The expansion of the universe means that all galaxies are moving away from the center of the universe. If you lived at the center, you'd see all the galaxies moving away from you. Since this is what we see, we must be at the center. Some galaxies are at the edge of the expansion. They're moving into empty space, so if you lived in one of the galaxies at the edge of the universe, you'd see that you are moving into a region where there are no galaxies.*

Making Sense of the Universe and Expansion

5) Do you agree or disagree with this students' description about the center and edge of the expanding universe? Explain your reasoning.

One way to try to understand and envision the expansion of the Universe is by creating analogies that attempt to model the different aspects of our real expanding Universe. Keep in mind that not *all* properties of the analogy apply to the real universe - only some properties do. One way to model the expanding Universe is to use a "balloon" analogy. In this analogy, the space and time of the universe are modeled by the "surface" or "skin" of an expanding balloon. Galaxies only exist on the surface of the balloon. Light can only travel on the surface of the balloon.



6) Do objects, light, or events in the universe also exist inside or outside of the balloon's surface in this analogy?

7) If you were to travel from galaxy to galaxy along the surface of the balloon, would you ever encounter an edge?

8) If you were to travel across the entire surface of the balloon universe, would you ever find a location that you would consider to be the center of the universe?

Making Sense of the Universe and Expansion

9) Consider the following debate between two students about their answers to the previous questions:

Student 1: *Someone living on the surface of this balloon universe will definitely encounter an edge and a center. All they have to do is look from their location across the inside of the balloon to a location on the other side. The center of the inside of the balloon is the center of the universe, and the far side would be the edge of what they could see. So there's definitely a center and an edge to the universe in the balloon analogy.*

Student 2: *I think you misunderstand the analogy. The surface of the balloon is supposed to be the entire universe. You can't look through the inside of the balloon to the other side so there is no center in the middle or edge on the other side. In this analogy, people living in the balloon universe would never see a center or an edge.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

10) In this analogy, do galaxies move relative to one another because they are traveling across the surface of the balloon, or do they move relative to one another because the balloon is expanding?

11) Imagine you lived in a galaxy on the surface of the balloon. As the balloon expands, would all the other galaxies appear to move toward you or away from you?

12) Would your answers to the previous question be the same regardless of the galaxy in which you live, or would it change depending on the galaxy you inhabit?

Making Sense of the Universe and Expansion

13) The balloon analogy is a helpful way to think about expansion, but no analogy is perfect. Some aspects of the real universe are captured by this analogy while others are not. Below are several properties of the real universe. For each, state whether it is accurately captured by the balloon analogy or not, and explain your reasoning.

a) The real universe has no center.

b) The real universe has no edge.

c) The real universe is expanding.

d) Light can travel in straight lines in the real universe.

e) The real universe's expansion does not cause galaxies to change size.

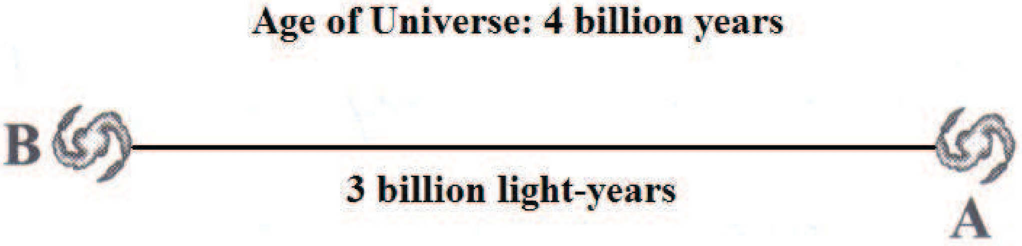
f) The real universe is 4-dimensional.

B.3 Expansion, Lookback Times, and Distances

The fall 2009 version of the “Expansion, Lookback Times, and Distances” lecture-tutorial begins on the next page.

Expansion, Lookback Times, and Distances

When the universe was 4 billion years old, Galaxy A was 3 billion light-years away from Galaxy B, as shown below. Imagine that the universe was not expanding, so the distance between Galaxy A and Galaxy B would not change over time.



1) A star explodes in Galaxy B producing a large amount of light. How long will the light from this explosion take to reach Galaxy A?

2) How far did the light travel on its journey to Galaxy A?

3) How old will the universe be by the time the light from the explosion reaches Galaxy A?

Because light takes time to travel from place to place in the universe, when we look at the night sky we are seeing stars and galaxies as they appeared in the past. For example, if we see a galaxy 1 million light-years away, we are seeing what the galaxy looked like 1 million years ago. We often use the term *lookback time* to describe how far back in time we are seeing. So if we see a galaxy as it appeared 1 million years ago, we say it has a lookback time of 1 million years.

4) What is the lookback time inhabitants of Galaxy A associate with Galaxy B when they see the light from the explosion?

The real universe is expanding. This means the distance between galaxies is constantly increasing. Imagine that Galaxy A and Galaxy B are in an expanding universe.

5) While the light from the explosion is traveling from Galaxy B to Galaxy A, does the distance between the two galaxies stay the same, become larger, or become smaller?

Expansion, Lookback Times, and Distances

6) By the time the light from the explosion reaches Galaxy A, is the distance to Galaxy B more than, less than, or exactly 3 billion light-years?

7) By the time the light from the explosion reaches Galaxy A, has more than, less than, or exactly 3 billion years elapsed since the star exploded?

8) By the time the light from the explosion reaches Galaxy A, will the total distance traveled by the light be more than, less than, or exactly 3 billion light-year?

9) When the inhabitants of Galaxy A see the light from the explosion, are they looking at an event with a lookback time of more than, less than, or exactly 3 billion years in the past?

10) Consider the discussion between two students regarding their ideas about two distant galaxies in an expanding universe.

Student 1: Let's say light takes 5 billion years to travel from one galaxy to another. This means the two galaxies were separated by 5 billion light-years when the light began its journey.

Student 2: If the light traveled for 5 billion years, then the distance between the two galaxies must have been less than 5 billion light-years when the light began its journey because the distances between galaxies are always increasing in the expanding universe.

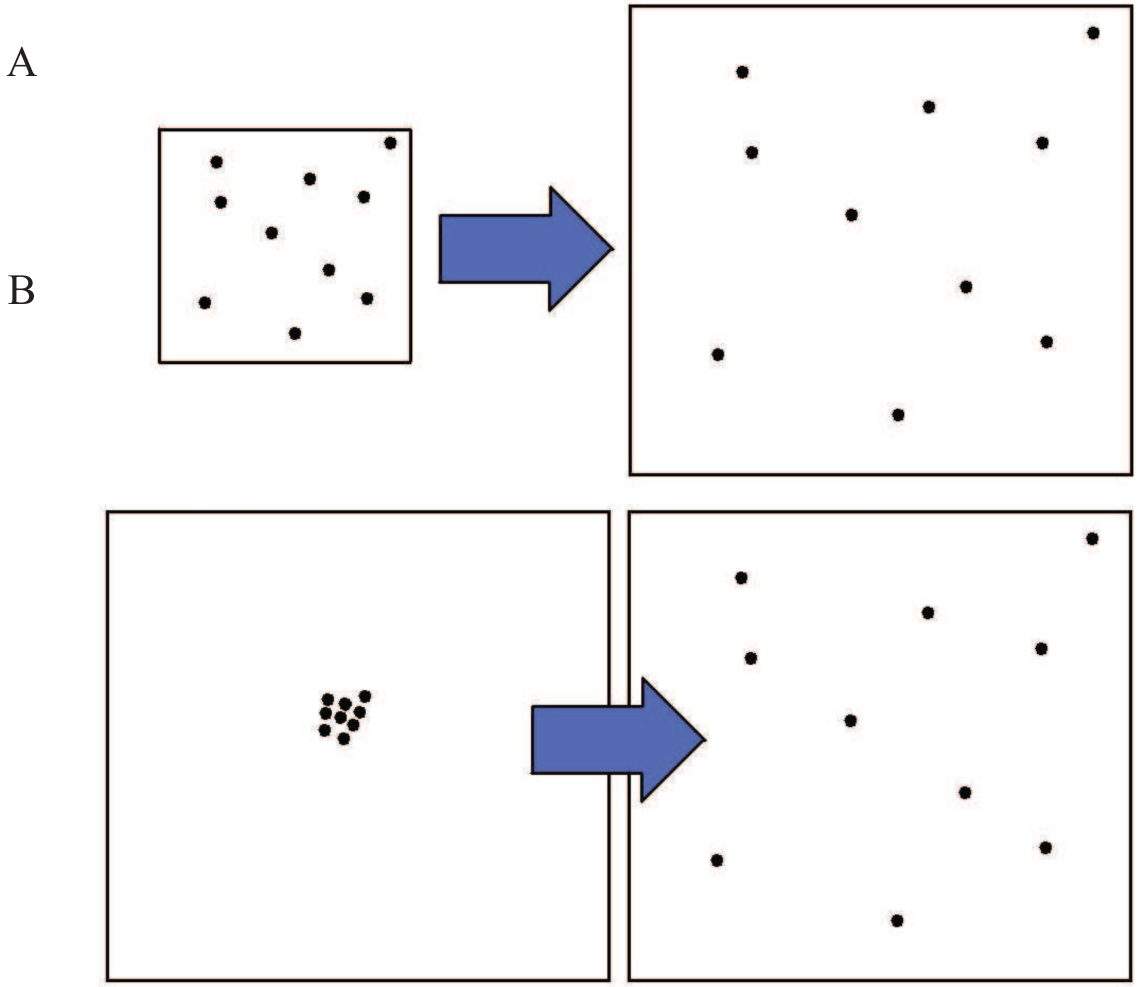
Do you agree with either or both of the students? Explain your reasoning.

B.4 The Big Bang

The fall 2009 version of “The Big Bang” lecture-tutorial begins on the next page.

The Big Bang

Consider the drawings (A and B) provided below which shows different ways of thinking about how the universe changes over time. The dots in each diagram represent pieces of matter.



1) Which drawing, A or B is a better representation of the universe we observe? Explain your reasoning.

The Big Bang

2) In Diagram A, is the universe becoming bigger, smaller, or staying the same size over time?

3) In Diagram B, is the universe becoming bigger, smaller, or staying the same size over time?

4) Two students are debating their answers to Questions 2 and 3:

Student 1: *Both diagrams show the universe becoming bigger. In Diagram A, the box has become larger. In Diagram B, the pieces of matter have spread out.*

Student 2: *I disagree. Only Diagram A shows the universe becoming bigger. In Diagram B the size of the box doesn't change. The pieces of matter are just moving into already exsiting empty space in a universe whose size doesn't change.*

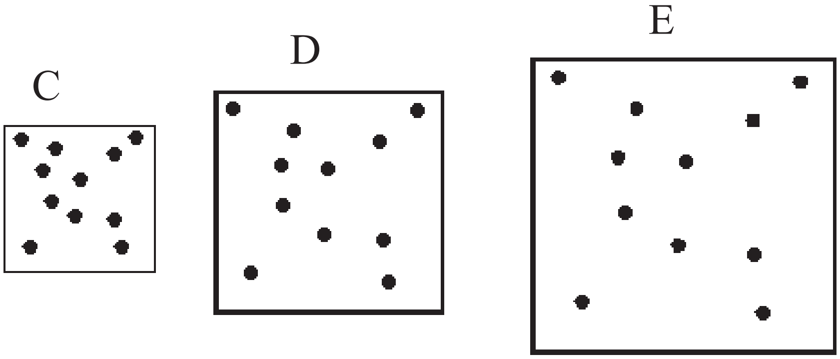
Do you agree or disagree with either or both of the students? Explain your reasoning.

5) Both drawings show the distance between matter increasing over time.
a) Which of the drawings shows this happening as the result of space expanding and which is a result of an outward explosion?

b) Which of these drawings is a correct representation of our universe? Is your answer to this question consistent to your answer to Question 1? Explain your reasoning.

The Big Bang

Consider the three drawings (C, D and E) shown below. These drawings each represent the same region of space, but at different times during the history of the universe.



6) Which drawing shows the region of space at the earliest point in the history of the universe? Explain your reasoning.

7) In which drawing does the region of space have:
   a) the highest density?

   b) the greatest concentration of energy?

   c) the highest temperature?

   Explain your reasoning.

8) Imagine you could watch the history of the universe like a movie playing backward. The movie starts today and ends at the beginning of the universe. Describe what you would see as the movie played and you looked further back in time. Your answer should discuss how the universe changes in terms of its temperature, density, and size.

The Big Bang

Your answers to the previous questions are all part of the *Big Bang Theory*. The Big Bang Theory does not say what the universe was like at the very first moment of time, which was about 13.7 billion years ago. It does, however, tell us how the universe changed after its first moment of existence.

9) Three students are discussing their understandings of the Big Bang Theory:

Student 1: *I think I understand the Big Bang now. At the beginning, all the matter in the universe was compacted into a small, hot, dense ball. This ball of matter then exploded into empty space. When we look at the universe, we see galaxies moving away from us. The Big Bang model explains this, since all matter should be flying away from the center point of the explosion.*

Student 2: *I disagree. I think what the Big Bang Theory is saying is that the entire universe was once really small, dense and hot and it expanded over time. There wasn't an explosion of matter into empty space. Instead, the universe carried galaxies and other matter away from each other as it expanded in size.*

Student 3: *You're both wrong. I agree that the universe was once smaller in size and that pieces of matter have been carried away from each other by the expansion of the universe. But remember how we learned from Einstein's equation* $E = mc^2$ *that matter can be converted into energy and energy can be converted into matter? I think this means that if we go back to the beginning of the universe, it would be so small, and its temperature so hot, that matter itself can't exist. I bet at the very beginning, the universe would have been infinitely small and composed of pure energy with no matter at all.*

State which students you agree or disagree with? Explain your reasoning.

The Big Bang

10) Based on your previous answers, complete the following sentences:

    The Big Bang Theory says that the universe started out with a very __________ temperature, a very ___________ density, and a very ___________ size. Originally, there was ______ matter, only pure ___________. From this initial state, the universe ____________ in size. This caused its temperature and density to ___________. When the temperature was cool enough, energy could transform into ___________.

11) Look at drawing A again. Next to drawing A, make a drawing of what you think the universe would have looked like at the very first instant it existed.

B.5 Dark Matter

The fall 2009 version of the “Dark Matter” lecture-tutorial begins on the next page.

Dark Matter

Part I: Motions of Planets

An object's orbit depends on the "mass inside" its orbit (also known as *interior mass*). For a planet in our Solar System, you can find the interior mass by adding the mass of the Sun to the mass of each object between the Sun and the planet's orbit. For example, the interior mass to Earth's orbit would be the mass of the Sun plus the mass of Mercury plus the mass of Venus.

Here is a table that lists each planet, the mass inside each planet's orbit, and the speed at which the planets orbit the Sun.

Planet	Interior Mass (solar masses)	Orbital Speed (km/s)
Mercury	1	47.9
Venus	1.0000027	35.0
Earth	1.0000057	29.8
Mars	1.0000060	24.1
Jupiter	1.00096	13.1
Saturn	1.0012	9.66
Uranus	1.0013	6.81
Neptune	1.0013	5.43

1) Fill in the blanks to complete the following sentences. It may be helpful to base your responses on the information provided in the table above.

   There are ______ planets inside Neptune's orbit and ______ planets inside Mercury's orbit. The interior mass for Neptune’s orbit is ________________________ *(much greater than/approximately the same as/much less than)* the interior mass of Mercury’s orbit.

2) Where is the vast majority of mass in the solar system located? What object or objects account for most of this mass?

The mass inside a planet’s orbit determines how fast the planet moves because it affects the strength of the gravitational force felt by the planet. If there's roughly the same amount of mass inside the orbit of two planets, then the planet farther away will orbit slower because the gravitational force on the more distant planet is weaker.

Dark Matter

3) How does the gravitational force on planets farther from the Sun compare to the gravitational force on planets closer to the Sun? Explain your reasoning.

4) How does the orbital speed of planets farther from the Sun compare to the orbital speed of planets closer to the Sun? Is this consistent or inconsistent with all of your previous answers? Explain your reasoning.

If you could increase the amount of mass inside a planet's orbit, you would increase the gravitational force it would feel and thus increase the orbital speed of the planet. Imagine you were able to add a very, very large amount of mass distributed evenly *between* the orbits of Jupiter and Saturn.

5) Which planet(s) will experience an increase in gravitational force and an increase in orbital speed from this added mass? Explain your reasoning.

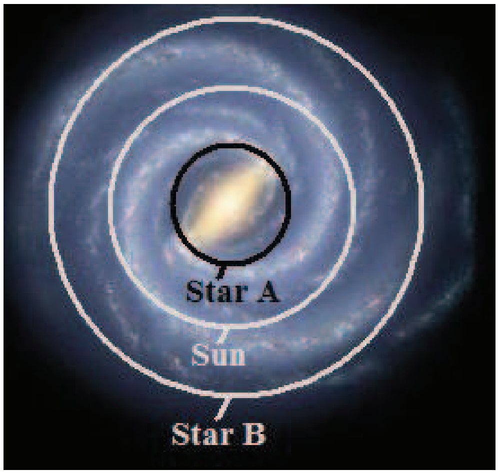
Part II: Motions of Stars

One way to estimate the amount of mass in a spiral galaxy is by looking at how much light the galaxy gives off. Where there is more light there must be more stars and hence more mass. When astronomers measure the amount of light coming from different parts of the galaxy, they find that stars are very concentrated in the central bulge of the galaxy and more spread out as you look outward through the disk.

Dark Matter

6) Based on the information provided above, where do you expect most the mass of a galaxy to be concentrated?

Here is a drawing of the Milky Way. The orbits of three stars are labeled. Star A is a star on the edge of the Milky Way's bulge. The Sun's orbit is shown at approximately the correct position. Star B is a star located farther out in the disk than the Sun.

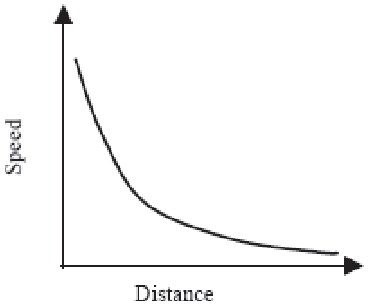


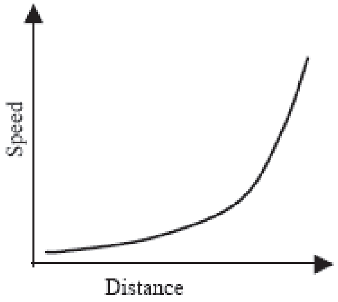
7) Based on your answer to question 6 and the location of each star from the center of the galaxy, rank the force of gravity on Star A, Star B, and the Sun from greatest to least. Explain your reasoning.

8) Based on your answer to Question 7, rank the orbital speeds of Star A, Star B, and the Sun from greatest to least. Explain your reasoning.

Dark Matter

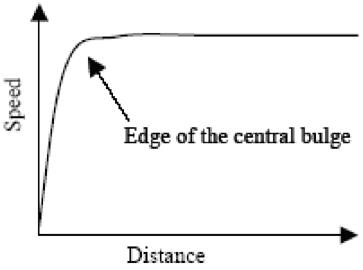
A graph of the orbital speed of stars versus their distance from the galaxy's center is called a *rotation curve*. Here are two possible rotation curves:





9) Which of the above rotation curves best represents the relationship you described in Question 8? Explain your reasoning.

Astronomers have measured how fast stars orbit around the center of our galaxy. Here is the real rotation curve for the Milky Way:



10) Based on their orbital distance from the center of the galaxy, make three dots on the above rotation curve to represent Star A, Star B, and the Sun. Be sure to label which mark belongs to each star.

11) Using this rotation curve, provide a new ranking for the orbital speed of Star A, Star B, and the Sun, from greatest to least. Describe any differences between this ranking and the one you provided in question 9.

Dark Matter

12) Based on your answer to Question 11, provide a new ranking for the gravitational force on Star A, Star B, and the Sun, from greatest to least. Explain your reasoning.

13) Based on your answers to question 10, 11 and 12, would you say that the mass of the Milky Way Galaxy is concentrated at its center (as is the case with our Solar System)? Explain your reasoning.

14) Two students are debating their answers to the previous questions:

Student 1: *Stars far from the center of the Milky Way are all moving at about the same speed. This means there must be more mass throughout the outer regions of the galaxy than we can see. If there were not, Star B would be moving slower than the Sun.*

Student 2: *I disagree. There are fewer stars in the outskirts of the Milky Way than in the center, so there's less mass out there than at the center. Most of the Milky Way's mass must be at its center. So I think the location of the mass doesn't affect how fast stars are moving.*

Do you agree with either or both of the students? Explain your reasoning.

Dark Matter

15) Astronomers initially thought there was more mass in the center of the galaxy than in the disk because there's more light coming from the galaxy's center than its disk. What observations have astronomers made that show these initial ideas are wrong? Explain your reasoning.

16) Is there more or less mass in the Milky Way's disk and halo than we can see? Explain your reasoning.

Appendix C

Fall 2009 Scoring Rubrics

The scoring rubrics for the fall 2009 begin on the next page.

Form A

Item 1a: Which graph or graphs show a universe that is expanding at a constant rate? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the expansion (or contraction) rate of the universe.
>
> **Correct answer:** F - This graph shows an expanding universe because galaxies that are farther away are moving away from us (positive velocity) faster than closer galaxies. Since the slope is constant, the rate at which the universe is expanding must be constant.

Item 1b: Which graph or graphs show a universe that is contracting at a constant rate? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the expansion (or contraction) rate of the universe.
>
> **Correct answer:** B - This graph shows a contracting universe because galaxies that are farther away are moving toward us (negative velocity) faster than closer galaxies. Since the slope is constant, the rate at which the universe is contracting must be constant.

Item 1c: Which graph or graphs show a universe that is expanding at a faster and faster rate over time? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the changing expansion (or contraction) rate of the universe, relative to lookback time.
>
> **Correct answer:** A - The rate at which the universe is expanding must be changing over time since the slope of the graph is changing with respect to distance. A universe expanding faster and faster over time must have a slope that gets steeper over time. This means the slope must be flatter at large distances since we are looking further back in time as we look farther away in distance.

Item 1d: Which graph or graphs show a universe that is expanding at a slower and slower rate over time? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the changing expansion (or contraction) rate of the universe, relative to lookback time.
>
> **Correct answer:** D - The rate at which the universe is expanding must be changing over time since the slope of the graph is changing with respect to distance. A universe expanding slower and slower over time must have a slope that gets flatter over time. This means the slope must be steeper at large distances since we are looking further back in time as we look farther away in distance.

The scoring rubric for these four items is on the next page.

Scoring rubric for items 1-4.

Overall Score	Answer Choice	Reasoning Categories
3 - Student gives the correct answer with the correct reasoning.	A - Student selects graph A.	I - Description includes all or part of the following statement: "† is ± and is * at a ‡." *See below for a complete explanation of this reasoning category and symbols.*
	B - Student selects graph B.	
2 - Student gives the correct answer with incorrect or incomplete reasoning.	C - Student selects graph C.	II - Student talks about how when we look deep into space we are looking back in time.
	D - Student selects graph D.	III - Student says her/his reasoning is opposite his/her reasoning on the previous item.
1 - Student gives an incorrect answer.	E - Student selects graph E.	IV - Student says s/he is guessing and/or unsure of why s/he chose that answer.
	F - Student selects graph F.	
0 - Student gives no answer, offers irrelevant information, or claims s/he has no idea how to answer.	G - Student selects graph G.	V - Student says s/he can't give an answer without more information.
	H - Student selects graph H.	VI - Student gives no reasons for answer.
	Z - Student selects no graph.	VII - Student gives some other reason not captured by categories I-X above.

Each response receives an overall score (0-3) as when as a letter (A-H or Z) to denote which graph the student chose. A Roman numeral (I-VII) is also assigned to each response to mark the student's reasoning. Reasoning Category I is further subdivided, as described on the next page.

Reasoning Category I explained - Many students justify their selections by noting that the graph or some aspect of the graph is 1) positive or negative (i.e. above or below the x-axis), 2) increasing, decreasing, constant, or curved, and 3) changing with an increasing, decreasing, or constant rate. Some students include all of these elements in the reasons they wrote; most only have a subset of these reasons. To better distinguish between these various possibilities, there are various subclasses for Reasoning Category I:

When the student writes "† is ± and is * at a ‡," append, when appropriate, the following symbols to I.

† = v *if the student discusses velocity or speed*

= d *if the student discusses distance*

= s *if the student discusses slope*

= r *if the student discusses rate*

= ℓ *if the student discusses the line*

= g *if the student discusses the graph*

= o *if the student discusses some other feature*

= i *if the student discusses an unspecified feature or "it"*

± = + *if the subject † is positive (above the x-axis)*

= - *if the subject † is negative (below the x-axis)*

* = INC *if the subject † is increasing*

= DEC *if the subject † is decreasing*

= CON *if the subject † is constant/straight*

= EXP *if the subject † is curved or exponentiating*

‡ = R1 *if the rate at which subject † changes increases*

= R2 *if the rate at which subject † changes decreases*

= R3 *if the rate at which subject † changes stays constant*

Example responses and scores:

A response to item 1a: *"F. The line is straight, making the expansion constant and it is increasing both Velocity and Distance."*

> **How this response is scored:** 3F_I_lCON_vINC_dINC. This student gave the correct answer (F) and a correct reason. She thus receives an overall score of 3. In her answer, she discusses how the line (l) is constant (CON), the velocity (v) is increasing (INC), and the distance (d) is increasing (INC).

A response to item 1b: *"B, graph constantly going down."*

How this response is scored: 2B_I_gDEC_R3. This student gives the correct answer (B) but with an incomplete reason, thus leading to an overall score of 2. In her answer, she says the graph (g) is going down (DEC) constantly (R3).

A response to item 1c: *"D"*

How this response is scored: 1D_VI. This student gave the wrong answer (D, hence leading to an overall score of 1) and gave no reason (VI).

A response to item 1d: *"A,B,E. Velocity is decreasing as the distance is increasing"*

How this response is scored: 1_ABE_vDEC_dINC. This person earns an overall score of 1 because he gave the wrong answer (A, B, and E). His reason talks about the velocity (v) decreasing (DEC) and the distance (d) is increasing (INC).

Item 2: Below are some plots (A - C) of how fast galaxies are moving versus their distance from us. Which of these three plots predicts an old universe? Explain your reasons for making your selection.

What this item measures: This item measures whether or not students can use Hubble plots to reason about the age of the universe.

Correct answer: C - This graph goes with the old universe because it has the smallest (flattest) slope, which means it corresponds to a universe with the slowest expansion rate, so it will take a long time for the universe to reach its current size.

The scoring rubric for item 2 is on the next page.

Scoring rubric for item 5.

Overall Score	Answer Choice	Reasoning Elements	
3 - Student gives the correct answer with correct and complete reasoning. A complete reason includes the following elements: 1) The graph has a flat slope which 2) means the expansion rate is slow and which means 3) the universe took along time to reach its current size. 2 - Student gives the correct answer with incorrect or incomplete reasoning. 1 - Student gives an incorrect answer. 0 - Student gives no answer, offers irrelevant information, or claims s/he has no idea how to answer.	A - Student selects graph A. B - Student selects graph B. C - Student selects graph C. Z - Student selects no graph.	I - Student talks about the time the universe took to reach the current size.	XXII - Student talks about "it" or something unspecified being constant.
		II - Student talks about large distances.	XXIII - Student talks about a large change in distance.
		III - Student talks about short distance.	XXIV - Student talks about a small change in distance.
		IV - Student talks about constant distance.	XXV - Student talks about a large change in velocity.
		V - Student talks about a fast velocity/speed.	XXVI - Student talks about a small change in velocity.
		VI - Student talks about a slow velocity/speed.	XXVII - Student talks about a large change in slope.
		VII - Student talks about a constant velocity.	XXVIII - Student talks about a small change in slope.
		VIII - Student talks about a large (steep) slope.	XXIX - Student talks about a large change in the expansion rate.
		IX - Student talks about a small (flat) slope.	XXX - Student talks about a small change in expansion rate.
		X - Student talks about a constant slope.	XXXI - Student talks about a large change in the graph.
		XI - Student talks about a high expansion rate.	XXXII - Student talks about a small change in the graph.
		XII - Student talks about a low expansion rate.	XXXIII - Student talks about a large change in the line.
		XIII - Student talks about a constant expansion rate.	XXXIV - Student talks about a small change in the line.
		XIV - Student talks about the graph being "high" or "large."	XXXV - Student says the expansion rate/velocity was high in the past and/or slow in the future.
		XV - Student talks about the graph being "low" or "small."	XXXVI - Student says the expansion rate/velocity was low in the past and/or high in the future.
		XVI - Student talks about the graph being constant.	
		XVII - Student talks about the line being "high" or "large."	XXXVII - Student says age is independent of distance and/or velocity.
		XVIII - Student talks about the line being "low" or "small."	XXXVIII - Student says there's not enough information.
		XIX - Student talks about the line being constant.	XXXIX - Student says s/he has no idea.

		XX - Student talks about "it" or something unspecified being "high" or large."	XXXX - Student gives no reason or leaves answer field blank.
		XXI - Student talks about "it" or something unspecified being "low" or small."	XXXXI - Student gives some other reason not listed above.

NB: You may assign multiple reasoning elements to a singe response.

Example responses and scores:

"C - The universe is because the slope of the graph is small, meaning it is moving/expanding slowly"

> **How this response is scored:** 2C_IX_XII. This student gives the correct answer (C), but for an incomplete reason. He talks about the slope being small (IX) and a slow expansion rate (XII).

"A. Because the velocity is still the same but the distance is very high."

> **How this response is scored:** 1A_VII_II. This student talks about a constant of velocity (VII) and a large distance (II). He gets an overall score of 1 since he gave the incorrect answer of A.

"I really don't know if any of them are old. If I had to guess, I would maybe say B because it seems to have developed a positive correlation b/w distance & velocity (maybe making it stable?)"

> **How this response is scored:** 1B_XXXIX_XXXXI. This student says she doesn't know (XXXIX), but guesses B which is incorrect. This gives her an overall score of 1. Her reasoning does not match any of the categories listed above (XXXXI).

Form B

Item 1: Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." Provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes students' conceptions of how the universe increases in size. Does a student have an expert-like conception, in which case s/he talks about how galaxies are getting farther from each other due to the stretching (expansion) of space (or spacetime)? Does the student incorrectly view the expansion as affecting the distances between all objects, including stars, planets, and other objects smaller than galaxies? Does the student incorrectly think that the "expansion of the universe" refers to growth in our knowledge or the creation of new objects/locations over time?

The scoring rubric for item 1 is on the next page.

Scoring rubric for item 1.

Score	Response Characteristics	Reasoning Elements
4	Students in category 4 meet the same criteria as students in category 3 EXCEPT they only discuss increasing distances and/or redshifts in the context of galaxies (and not planets, stars, or other objects smaller than galaxies). Students may be placed in category 4 without explicitly saying that • the universe has no center, • the universe has no edge, • the Big Bang refers to the evolution of the universe from a hot, dense state. If the student does discuss one of these elements and makes an incorrect statement then s/he cannot be placed in category 4.	I - Student says the size of the universe increases over time. II - Student says the universe has a center. III - Student says the universe has no center. IV - Student says the universe has an edge. V - Student says the universe has no edge. VI - Student talks about redshifts/Doppler shifts. VII - Student says space(time) is growing/stretching. VIII - Student talks about the movement of galaxies and/or their increasing distances.
3	Students in this category say that • The universe increases in size and/or • The distances between all objects increase. Additionally, they must make at least one of the following claims and not contradict the others: • The universe has no center. • The universe has no edge. • The Big Bang refers to the evolution of the universe from a hot, dense state.	IX - Student talks about the movement of stars and/or their increasing distances. X - Student talks about the movement of planets and/or their increasing distances. XI - Student talks about the movement of objects (something unspecified or not a star, planet, or galaxy) and/or their increasing distances. XII - Student says the distances between everything increase.
2	Students in this category say that the universe increases in size and either provide no other information or claim one or more of the following: • The universe has a center. • The universe has an edge. • The Big Bang is an explosion. • The distances between all objects increase.	XIII - Student says farther objects move away faster. XIV - Student talks about the Big Bang. XV - Student talks about an explosion. XVI - Student says the early universe was once hot, small, and/or dense. XVII - Student says we learn more about the universe over time.
1	Students in this category describe the expansion of the universe as something other than the universe getting larger over time. The two most popular answers that fit in this category are • Expansion refers to learning more about the universe over time. • Expansion refers to the creation of new objects over time.	XVIII - Student talks about how we are looking further back in time as we look farther into space. XIX - Student says new things are created in the universe over time. XX - Student gives irrelevant information. XXI- Student gives some other reason not specified above. XXII - Student has no idea.
0	Students in this category write nothing (the answer field is blank), or they provide information that doesn't answer the question, or they say they have no idea and provide no further information.	XXIII - Answer field is blank or the student provided no reason.

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response will contain every element in a given score category, you must also list the reasoning elements the student uses.

Example responses and scores:

"There was a man named 'hubble' who had a telescope and observed galaxies 'red shifting' or moving away. Now when he observed this he had the best telescope and plotted the rate of speed of expansion. He noticed that although the closer galaxies are red shifting at a higher rate, the most far away galaxies were moving away (Red shifting) at a much higher speed. From this observation astronomers deducted the universe is expanding from all far away galaxies are red shifting At a higher speed"

> **How this response is scored:** 4_VI_VIII_XIII. This student talked about redshifts (VI), how distances between galaxies are increasing (VIII), and how the farther away galaxies are, the faster they move away from us (XIII). Since the student talked about distances to galaxies increasing and did not make any incorrect statements about expansion, he receives an overall score of 4.

"'the universe is expanding' because as time goes on, matter is moving further and further away from other matter. Temperature and density have decreased over time as the universe is expanding and matter has moved away"

> **How this response is scored:** 3_XI_XVI. This student talks about otherwise unspecified pieces of matter moving away from one another (XI) and talks about how the universe used to be hotter and dense (XVI). While none of these pieces of information are necessarily wrong, she only gets an overall score of 3 because she does not specify whether expansion only affects the distances between galaxies or whether the distances between smaller objects are also affected.

"There was a big bang which exploded out everything in the universe. The leading edge of this bang has been expanding ever since."

> **How this response is scored:** 2_XIV_XV_XII_IV. This student talks about the Big Bang (XIV) as an explosion (XV). His answer indicates that distances between everything are increasing (XII) and he mentions an edge (IV). These reasoning elements give him an overall score of 2.

"I don't think that it is acculuy expanding in a physical sense, but instead or knowledge of the univers and the areas that we have discovered is expanding with an increase in technology and invesments in sciens"

> **How this response is scored:** 1_XVII. This student denies that the universe is physically growing. Instead, he says that expansion refers to our increase in knowledge over time (XVII), placing him in category 1.

Item 2: Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory." Provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether students think of the Big Bang as the beginning of the expansion of the universe, NOT an explosion of pre-existing matter into empty space, nor as the origin of something smaller than the universe (the Earth, Solar System, Galaxy, etc.).

The scoring rubric for item 2 is on the next page.

Scoring rubric for item 2.

Score	Response Characteristics	Reasoning Elements
3	Students in category 3 talk about the Big Bang using at least one of the following elements: • It's the beginning of the universe, space, and/or time. • It's when (some) of the elements were created. • It refers to how the universe was once hot, small, and dense. Additionally, they also include at least one of these elements (and do not contradict any of the others): • It was not an explosion. • It was the beginning of expansion. • Matter formed from energy. • It encompassed all of the universe.	I - Student says the Big Bang is the beginning of the universe. II - Student says the Big Bang is the beginning of expansion. III - Student says the Big Bang was the beginning of something smaller than the universe. IV - Student says the Big Bang is an event that happened to something smaller than the universe. V - Student says the Big Bang is the beginning of space. VI - Student says the Big Bang is the beginning of time. VII - Student talks about the creation/production of elements.
2	Students in category 2 talk about the Big Bang using at least one of the following elements: • It's the beginning of the universe, space, and/or time. • It's when (some) of the elements were created. • It refers to how the universe was once hot, small, and dense. Students in this category either provide no further information or they use one of the following elements: • The Big Bang was an explosion. • Matter existed before the Big Bang. • The Big Bang was an event that happened in empty space. Additionally, any student whose ideas about the universe do not fit into categories 3 or 1 should be placed here.	VIII - Student says the Big Bang was an explosion. IX - Student says the Big Bang was not an explosion. X - Student says matter existed before the Big Bang. XI - Student says there was a dense piece of matter before the Big Bang. XII - Student talks about matter coming together before the Big Bang. XIII - Student says matter formed from energy. XIV - Student says the early universe was hot, dense, and/or small. XV - Student says the Big Bang was an event that happened in empty space.
1	Students in category 1 talk about the Big Bang as the beginning of something smaller than the universe or and event that happened to something smaller than the universe.	XVI - Student gives irrelevant information. XVII - Student gives some other reason not specified above. XVIII - Student says s/he has no idea.
0	Students in category 0 write nothing (the answer field is blank), or they provide information that doesn't answer the question, or they say they have no idea and provide no further information.	XIX - Answer field is blank or the student provided no reason.

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response with contain every element in a given score category, you

must also list the reasoning elements the student uses.

Example responses and scores:

"The Big Bang Theory, in a nutshell, says that the universe started out as very hot + dense, and over time that energy was converted into matter as the distance between matter increased and as the universe expanded."

> **How this response is scored:** 3_XIV_XIII_II. This student gets and overall score of 3 because in her response she mentioned expansion (II), the formation of matter from energy (XIII), and how the universe was once hot and dense (XIV).

"Big Bang Theory (not the TV show) means that all the stuff, matter, energy, whatever that exists in the universe now was once scrunched up into a teeny tiny little thing which then exploded and expanded rapidly."

> **How this response is scored:** 2_X_XI_II_VIII. This student receives a 2 for this response because she said the Big Bang is the beginning of expansion (II), a dense piece of matter existed before the Big Bang (X and XI), and the Big Bang was an explosion (VIII).

"'the big bang theory' is when an asteroid that was headed toward earth struck the earth and every thing that was alive died --- then as time went on things started growing and living again."

> **How this response is scored:** 1_IV. This response earns an overall score of 1 because it talks about the Big Bang as an event that happened to something smaller than the universe (IV).

"Big Bang theory means an idea or a way of critical thinking between astronomers, and/or scientists to the way they see the earth and universe through telescope or even without one through basic observation."

> **How this response is scored:** 0_XVI. This response does not really answer the question and gives a lot of irrelevant information (XVI).

Item 3 (AZ version): Imagine you had a spaceship that could take you back in time to the time of the Big Bang. Would there have been any locations in the universe from which you could have watched the Big Bang from a distance? If yes, describe what it would have looked like and draw a picture. If no, explain why not.

Item 3 (CU version): If you were alive at the time of the Big Bang, would there have been any locations in the universe from which you could have watched the Big Bang from a distance? If yes, describe what it would have looked like and draw a picture. If no, explain why not.

> **What this item measures:** This item probes whether students think the Big Bang was an event that happened in a certain location in the universe or whether it was an event that encompasses the entire universe.

The scoring rubric for item 3 is on the next page.

Scoring rubric for item 3.

Score	Response Characteristics	Reasoning Elements
3	Students in category 3 have answers that say there was no space outside of the Big Bang. Additionally, they have no incorrect statements about the Big Bang.	I - Student talks about how the universe used to be small and/or how there was no space in the universe outside the Big Bang.
2	Students in category 2 have answers that fall into one of the following: • There was space outside of the Big Bang. • Answer is unclear as to whether or not there was space outside of the Big Bang. • Answer has incorrect ideas about the Big Bang (e.g. it was an explosion, matter existed before the Big Bang, etc.) but does not claim the Big Bang was something smaller than the beginning of the universe.	II - Student says there was space in the universe outside the Big Bang. III - Student says there was no time before the Big Bang. IV - Student says there was time before the Big Bang. V - Student says there was no matter before the Big Bang. VI - Student says there was matter before the Big Bang.
1	Students in category 1 have answers that say the Big Bang was something smaller than the beginning of the universe.	VII - Student says the universe began at the Big Bang.
0	Students in category 0 write nothing (the answer field is blank), or they provide information that doesn't answer the question, or they say they have no idea and provide no further information.	VIII - Students says there was nothing/no locations before the Big Bang.
		IX - Student says the Big Bang was an explosion. X - Student says the Big Bang wasn't safe. XI - Student talks about high temperatures. XII - Student talks about high density. XIII - Student talks about debris. XIV - Student says the Big Bang would not have been visible to the naked eye. XV - Student says the Big Bang would have been too hard to find. XVI - Student says the Big Bang would have taken too long for someone to see. XVII - Student says the Big Bang would have been too far away for someone to see. XVIII - Student says there was originally just energy. XIX - Student says you could have seen the Big Bang from somewhere in the Solar System. XX - Student says you could see the Big Bang from elsewhere in the galaxy/another solar system or star. XXI - Student says you could see the Big Bang from outside the Milky Way/from another galaxy.

	XXII - Student says you could see the Big Bang from some not otherwise specified location. XXIII- Student gives irrelevant information. XXIV - Student gives some other reason not specified above. XXV - Student says s/he has no idea. XXVI - Answer field is blank or the student provided no reason.

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response with contain every element in a given score category, you must also list the reasoning elements the student uses.

Example responses and scores:

"No because before the big bang there was no Universe + therefore no place in the Universe to go to see it."

> **How this response is scored:** 3_VII_I. This response receives an overall score of 3 because it says there was no place in the universe outside the Big Bang (I) and the universe began at the Big Bang (VII).

"no, since the big bang was the origin of the universe, and as such nothing really existed before it + matter only existed after that first instant."

> **How this response is scored:** 2_VII_VIII. This student talks about how the universe began at the Big Bang (VII) and how there was nothing before the Big Bang (VIII). However, he does not specify whether or not he thinks the Big Bang was surrounded by empty space, so he receives an overall score of 2.

"No because everything in the universe is made from the Big Bang it would be so big that the radiation would kill you therefore I guess you could from anywhere"

> **How this response is scored:** 2_II_X_VIII. This answer claims that there was space in the universe outside the Big Bang (II). He also says there was nothing before the Big Bang (VIII) and he expresses concern over the safety of this expedition (X). His overall score is a 2.

"You could have watched the Big Bang far away in the galaxy. The Big Bang was the formation of earth so therefore it could be visible from elsewhere"

> **How this response is scored:** 1_XX. This student receives a 1 because her answer claims that you could see the Big Bang from elsewhere in the galaxy (XX) because it was the formation of the Earth.

Item 4: If you could travel to any location in the universe, could you ever see the center of the universe? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes whether or not students think the universe has a center and why.

Scoring rubric for item 4.

Score	Response Characteristics	Reasoning Elements
3	- rejection of the idea of a center based on claim that universe is infinitely big/has no edges. - rejection of the idea of a center based on the idea that the universe looks the same no matter where you are.	I - Student talks about how the universe is infinite/has no edges. II - Student reasons using the idea of a center based on the idea that the universe looks the same no matter where you are.
2	- rejection of the idea of a center but with no justification or for reason other than those listed in category 3.	III - Student says the Sun is in the center of the universe.
1	- statement that the universe has a center or statement is unclear as to whether or not there is a center. - claim that we can't see the center due to our own ignorance about the universe or limitations in our technology. - claim that the center changes due to the fact that the universe is expanding.	IV - Student says the center is where the Big Bang happened/where everything is moving away from. V - Student says our ignorance about the universe prevents us from seeing its center. VI - Student says the limitations of our technology prevent us from seeing its center.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	VII - Student says there's no center or center of the universe changes because the universe is expanding and/or things are in motion. VIII - Student says the universe has no center because it does not have a shape/has an irregular shape. IX- Student gives irrelevant information. X - Student gives other reason not specified above. XI - Student says s/he has no idea. XII - Response field left blank or no reason given for answer.

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response with contain every element in a given score category, you must also list the reasoning elements the student uses.

Example responses and scores:

"There is not a center! wherever you are in the universe everything will appear to be moving away from you at an increasing rate. There is no edge, top or bottom."

> **How this response is scored:** 3_I_II. This score earns a 3 because the student says the universe has no edge (I) and that it looks the same no matter where you are (II).

"no, b/c its always expanding"

> **How this response is scored:** 2_VII. While the student correctly says no, he ties his answer to the fact that the universe is expanding (VII) with no further explanation. This is why he receives an overall score of 2.

"No, because we do not know where the center is, we can only see so much."

> **How this response is scored:** 1_V. This student receives a 1 because he does not deny the existence of the center. Instead, he says we cannot see it due to our lack of knowledge (V).

Item 5: If you could travel to any location in the universe, could you go to a place where there would be no galaxies in front of you? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes whether students think there is an edge to the universe (in the sense that there is an end to the distribution of galaxies).

Scoring rubric for item 5.

Score	Response Includes:
3	- no, there will always be galaxies around you
2	- yes, if you go to a black hole or some other unintended or not quite correct reason.
1	- yes, there are regions where distribution of galaxies peters out/where there are no galaxies
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.

NB: There are no reasoning elements listed since each score is associated with its own unique reasoning element.

Example responses and scores:

"Um...I suppose you could go to a black hole and stare into it. Black holes are really super dense and no light escapes them so you wouldn't be able to see galaxies."

> **How this response is scored:** This student gets a score of 2 because she talks about looking into a black hole.

"No, there are galaxies everywhere."

> **How this response is scored:** This response falls into category 3 since the student says there will always be galaxies around you.

"Yes that would be toward the outside of the universe."

> **How this response is scored:** This response earns a score of 1 because the student says there are regions with no galaxies.

Item 6: Which of the following is true? There may be more than one correct answer. Select all that apply.

a) Over time, the distances between widely separated planets in the solar system **always increase** and **never decrease**.
b) Over time, the distances between widely separated stars in the galaxy **always increase** and **never decrease**.
c) Over time, the distances between widely separated galaxies in the universe **always increase** and **never decrease**.
Explain your reasoning for your choice(s).

> **What this item measures:** This item tests whether or not students realize that the expansion of the universe only affects distances between galaxies, not distances between stars within a galaxy or planets within a solar system.

Scoring rubric for item 6.

Score	Response Includes:
2	- idea that expansion only affects the distances between galaxies.
1	- idea that expansion affects distances other than those on galactic scales or that we cannot tell.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.

NB: There are no reasoning elements listed since each score is associated with its own unique reasoning element.

Example responses and scores:

"C. Stars and planets move toward and away from each other in the universe."

> **How this response is scored:** This student gets an overall score of 2 because she says that expansion only affects the distances between galaxies.

"A, B. Planets + stars expand their orbits. I'm not sure if galaxies expand."

> **How this response is scored:** This student gets an overall score of 1 because she claims that the distances other than those between galaxies increase due to expansion.

Form C

Item 1: Galaxy X and Galaxy Y are currently 8 billion light-years apart in the expanding universe. How long will light from Galaxy X take to reach Galaxy Y?

a) less than 8 billion years b) exactly 8 billion years
c) more than 8 billion years d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

What this item measures: This item measures whether or not students realize the expansion of the universe affects the light travel time between two galaxies.

Correct answer: C - Normally, light takes 8 billion years to travel 8 billion light-years. However, since the universe is expanding, the distance between Galaxies X and Y is increasing while the light is traveling. This means the light will end up traveling more than 8 billion light-years by the time it reaches Galaxy Y, which means it will take more than 8 billion years to get there.

Scoring rubric for item 1.

Score	Response Includes	Answer Choice	Reasoning Elements
3	- right answer (C) - correct reason	A - Student choose answer A.	I - Student talks about the expanding universe/galaxies getting farther apart.
2	- right answer (C) - no reasons given or incomplete or incorrect reasons		II - Student talks about galaxies getting closer together.
		B - Student chooses answer B.	III - Student says light needs 8 billion years to travel 8 billion light-years.
1	- incorrect answer(s)		IV - Student says light-years are shorter than normal years.
0	- nothing (answer field left blank - information that does not answer the question - no idea	C - Student chooses answer C.	V - Student says light-years are longer than normal years.
		D - Student chooses answer D.	VI - Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.
		Z - Student makes no selection.	VII - Student says light may be slowed by something between the two galaxies.
			VIII - Student says there's not enough information.
			IX - Student gives irrelevant information.
			X - Student gives other reason not specified above.
			XI - Student says s/he has no idea.
			XII - Response field left blank or no reason given.

NB: You may assign multiple reasoning elements to a single response.

Example responses and scores:

"C, If the universe was stationary, it would take 8 billion years. Since the universe is expanding + we don't know at what rate, it is more than 8 billion."

> **How this response is scored:** 3C_I. This student gave the correct answer (C) and for the right reason since he talked about the effects of the expansion of the universe (I). Therefore, he gets an overall score of 3.

"C. Because the Galaxy is always expanding day by day and we have noticed if it expanded. And because light years are longer than years. They are years at the speed of light."

> **How this response is scored:** 2C_I_V. This student gives the right answer (C) and talks about the expansion of the universe (I). However, he also says light-years are longer than longer years (V). The student thus receives an overall score of 2.

"B. If they are 8 billion light years apart it should take 8 billion years for light to travel. But I'm not sure. I am have an immensely difficult time in regard to the wording of the question just asked me (light years - years, traveling light - that stuff)"

> **How this response is scored:** 1B_III. This student gets a 1 because he gave the wrong answer (B). His reasoning including the idea that light needs 8 billion years to travel between two galaxies separated by 8 billion light-years (III).

"A. less than because it can't take as long as how far it is apart. It doesn't make sense that it would take as long."

> **How this response is scored:** 1A_X. This student gave a wrong answer (A) so she gets an overall score of 1. Her reason did not fall into one of the defined categories (X).

Item 2: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize the temperature of the universe has cooled over time.

The scoring rubric for item 2 is on the next page.

Scoring rubric for item 2.

Score	Response Includes	Reasoning Elements
3	- claim that universe has cooled over time. - reason based on the expansion of the universe.	I - Student says the temperature increased. II - Student says the temperature decreased. III - Student says the temperature stayed the same.
2	- claim that universe has cooled over time. - no reasons given or incomplete or incorrect reasons.	IV - Student says the temperature changed, but doesn't specify if it went up or down. V - Student talks about expansion and/or how the density of the universe changed.
1	- any claim other than the universe has cooled over time.	VI - Student talks about the birth/formation of stars/a star.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	VII - Student talks about the birth/formation of planets/a planet. VIII - Student talks about the birth/formation of galaxies/a galaxy. IX - Student talks about the birth/formation of unspecified objects. X - Student talks about the death of stars. XI - Student talks about the death of planets. XII - Student talks about the death of galaxies. XIII - Student talks about the death of unspecified objects. XIV - Student talks about changes during the lives of stars/a star. XV - Student talks about changes during the lives of planets/a planet. XVI - Student talks about changes during the lives of galaxies/a galaxy. XVII - Student talks about changes during the lives of unspecified objects. XVIII - Student talks about how the universe is big. XIX - Student talks about competing effects canceling out. XX - Student gives irrelevant information. XXI - Student gives other reason not specified above. XXII - Student says s/he has no idea. XXIII - Response field left blank.

NB**:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"The temp. has gotten cooler. Again assuming no matter is created in the expansion process of the universe than while the universe has expanded it has become less dense and as you become less dense you cool off."

> **How this response is scored:** 3_II_V. This student gets a 3 because he says the temperature decreases (II) and he explains this in terms of the expansion of the universe (V).

"It could be getting colder as time goes on because as the universe expands there is more and more area for the sun and other giant stars to heat, and also the sun is continuously collapsing on itself (I think?) which might mean its getting smaller, just like other giant stars in the universe are which could also be making it cooler as time progresses."

> **How this response is scored:** 2_V_XVIII_XIV_II. This student correct says that the universe is cooling (II) and relates this to its expansion (V). However, she also relates the cooling to changes during the lives of stars (XIV) and mentions the big region (XVIII) stars have to heat. She thus receives an overall score of 2.

"It has to have changed b/c Galaxies are colliding, New stars are formed, and many stars die. The temperature can vary."

> **How this response is scored:** 1_IV_XVI_VI_X. This student says the temperature changed, but does not specify if he thinks it went up or down (IV) so he gets an overall score of 1. His answer also contains information on how galaxies change (XVI), the formation of stars (VI), and the death of stars.

Item 3: How does the <u>total amount of matter</u> in the universe right now compare to the total amount of matter in the universe at the very beginning of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize that the total amount of matter is not conserved in the universe and that the early universe was pure energy.

Scoring rubric for item 3.

Score	Response Includes	Reasoning Elements
3	- claim that there is more matter in the universe now. - reason based on the idea that the early universe was pure energy and/or contained no matter until the temperature dropped.	I - Student says there was no matter/only energy in the beginning. II - Student talks about the temperature cooling and/or that matter formed from energy. III - Student says there is more matter now because the universe is expanding. IV - Student says the amount of matter increases as objects form and/or evolve. V - Student says the amount of matter decreases as objects form and/or evolve. VI - Student says the amount of matter increases as objects interact and/or die. VII - Student says the amount of matter decreases as objects interact and/or die. VIII- Student says the amount of matter does not change. IX - Student gives irrelevant information. X - Student gives other reason not specified above. XI - Student says s/he has no idea. XII - Response field left blank or no reason given.
2	- claim that there is more matter in the universe now. - no reasons given or incomplete or incorrect reasons.	
1	- any claim other than there is more matter in the universe now.	
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	

***NB*:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"There is tons more matter now. Then there was tons of energy, but now it has transformed into matter."

> **How this response is scored:** 3_I_II. This student says there used to be a lot of energy in the universe (I) but that some of it was subsequently transformed into matter (II), so she gets an overall score of 3.

"I think there is more because the universe is expanding and things are being created along w/expansion to 'fill it up'"

> **How this response is scored:** 2_III_IV. This student says the amount of matter in the universe has increased, but explains this in terms of the expansion of the universe (III) and new objects forming (IV). Therefore, she gets an overall score of 2.

"There is just as much matter now then then. It has just expanded."

> **How this response is scored:** 1_VIII. This student gets a 1 because he says the amount of matter hasn't changed (VIII).

Item 4: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize that the density of matter has decreased over time due to expansion..

Scoring rubric for Item 4.

Score	Response Includes	Reasoning Elements
3	- claim that the density has decreased over time - reason based on the universe becoming bigger over time	I - Student says the density increases. II - Student says the density decreases. III - Student says the density is constant. IV - Student says the density changes, but doesn't specify how. V - Students says the amount of matter in the universe has changed. VI - Student says the amount of matter in the universe has not changed. VII - Student says the size of the universe has changed. VIII - Student says the size of the universe has not changed. IX - Student talks about objects forming over time. X - Student talks about objects evolving over time. XI - Student talks about objects dying/breaking up. XII - Student talks about the effects of gravity. XIII - Student talks about matter changing forms. XIV - Student gives some other reason not specified above. XV - Student gives irrelevant information. XVI - Student says s/he has no idea. XVII - Answer field is blank or no response given.
2	- claim that the the density has decreased over time - no reasons given or incomplete or incorrect reasons	
1	- any claim other than the density has decreased over time	
0	- nothing (answer field left blank) - information that does not answer the question - no idea	

***NB*:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"It was more dense @ the beginning, and is now less dense b/c the universe has expanded."

> **How this response is scored:** 3_II_VII. This student gets an overall score of 3 because she correctly said that the density has dropped (II) due to the expansion of the universe (VII).

"I think it's got less dense I mean there was none, then the Big Bang and now it formed."

> **How this response is scored:** 2_II_V_IX. This student gets an overall score of II because, although he said the density decreased (II), he also talked about how the amount of matter has changed (V) and objects forming (IX).

"It has changed, I couldn't say if its gotten larger or smaller but since D=M/V, and the mass has not changed, but I'm pretty sure V has, the D would be different."

> **How this response is scored:** 1_IV_VI_VII. This student says the density changes, but can't say how (IV), so she gets an overall score of 1. She also talks about how the amount of matter has not changed (VI), how the size of the universe has changed (VII).

Form D

Item 1: Which one graph best represents how objects actually move in our galaxy? ***Explain the reasoning behind your choice.***

> **What this item measures:** This item measures whether or not students can identify the correct rotation curve for the galaxy (Graph 2).

Scoring rubric for item 1.

Score	Response
2	- student selects graph 2
1	- student selects any graph other than graph 2
0	- nothing (answer field left blank) - information that does not answer the question - no idea

NB: Although the question asks for students to explain their reasoning, I did not find any patterns useful to highlight in the responses students gave. Most simply described the graph they chose. In the post-surveys, many mentioned "dark matter" but not at a level beyond declarative knowledge. This is probably okay, since subsequent items probe the students' understandings of rotation curves further.

Item 2: Based on your selection in Question 1, describe how the speeds at which stars A, B, and C orbit the galaxy compare to one another.

> **What this item measures:** This item measures whether students can use the rotation curve they selected in item 1 to correctly compare the speeds of the three stars.

Scoring rubric for item 2.

Score	Response
2	- student correctly relates the speeds of the three stars based on the graph s/he chose
1	- student incorrectly relates the speeds of the three stars based on the graph s/he chose
0	- nothing (answer field left blank) - information that does not answer the question - no idea

Even though item 3 is not "officially" scored, still record a score of 0, 1, or 2. A student earns a 2 if s/he ranks the forces based on the idea that the velocity of a star directly corresponds to the gravitational force it feels (e.g. a student who says all stars move with the same velocity should say that all stars feel the same force, or a student who says stars closer to the galaxy's center are faster should likewise assign larger forces to stars closer to the center). Otherwise, any other reason is scored a 1. If there is no answer or the student say s/he has no idea with no further comment, give her/him a 0.

Item 4: Based on your previous answers, what does this graph tell us about how matter is distributed in our galaxy? Explain your reasoning.

> **What this item measures:** This item measures whether students realize that matter is not concentrated in the center of the galaxy.

Scoring rubric for item 4.

Score	Response
2	- if the student indicated in her/his previous answers that s/he thinks the speeds decrease with increasing distance, then s/he answers this item by saying the matter in the galaxy must be concentrated in the galaxy's center. - if the student indicated in her/his previous answers that s/he thinks the speeds increase OR stay constant with decreasing distance, then s/he answers this item by saying the matter in the galaxy must NOT be concentrated in the galaxy's center.
1	- student gives an answer contrary to her/his previous answers - student does not explain whether matter is concentrated in the galaxy's center or not
0	- nothing (answer field left blank) - information that does not answer the question - no idea

Example responses and scores:

NB: All of these responses are taken from the same student.

Item 1: *"3. Objects closest to the center of the galaxy move fastest because they are receiving the most gravity and have to move fast enough so that they don't get pulled to the center"*

> **How this response is scored:** The student chose the wrong rotation curve, so he receives a 1.

Item 2: *"Star A is moving the fastest, then B, then C. Again, the closer the object to the center where gravity is pulling, the objects must be fast to avoid being pulled in."*

> **How this response is scored:** This answer is consistent with the graph the student selected in item 1, so he receives a 2 on this question.

Item 3: *"The net g force on A would be greatest because it is closest to the center. star B would be next. then Star C. The net gravitational force is dependent on its distance from the center"*

> **How this response is scored:** This answer is consistent with the student's previous answers, so he receives a 2.

Item 4: *"The majority of matter is located at the center of the galaxy where the gravitational force is strongest due to the high mass of objects such as a black hole?"*

> **How this response is scored:** This answer is consistent with the student's previous answers, so he receives a 2.

Appendix D

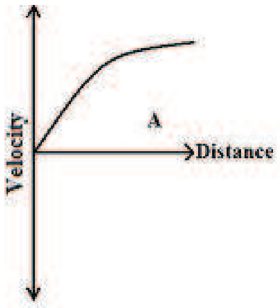
Spring 2010 Surveys

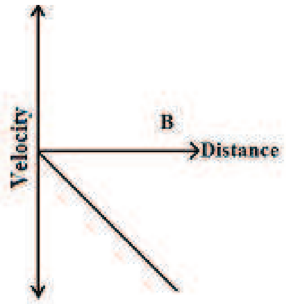
D.1 Form A

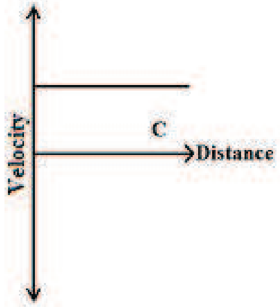
The spring 2010 version of Form A begins on the next page.

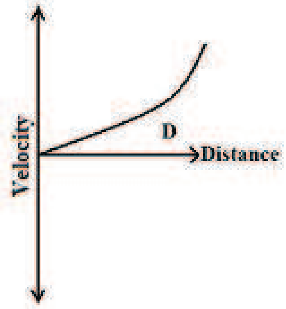
Name: ________________________________ **Conceptual Cosmology Survey - Form A**

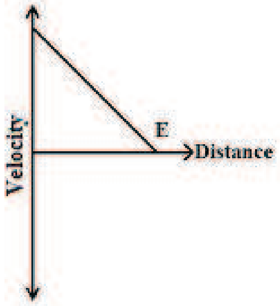
Below are eight graphs (A - H) showing how fast galaxies are moving (velocity) versus their distances away from us.

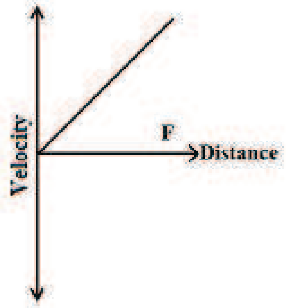


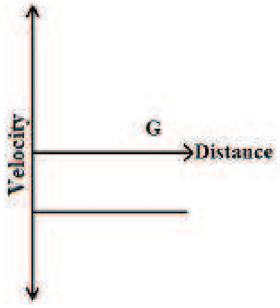


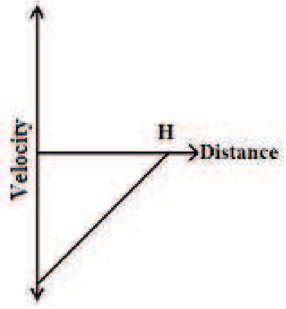












1) Which graph or graphs, if any, show a universe that is expanding at a constant rate? Explain your reasoning for your selection(s). If your answer is "none," explain why.

2) Which graph or graphs, if any, show a universe that is contracting at a constant rate? Explain your reasoning for your selection(s). If your answer is "none," explain why.

3) Which graph or graphs, if any, show a universe that is expanding at a faster and faster rate over time? Explain your reasoning for your selection(s). If your answer is "none," explain why.

4) Which graph or graphs, if any, show a universe that is expanding at a slower and slower rate over time? Explain your reasoning for your selection(s). If your answer is "none," explain why.

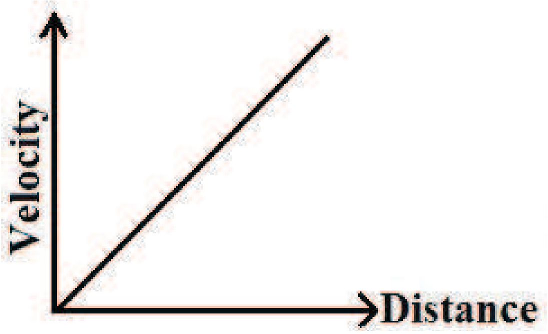
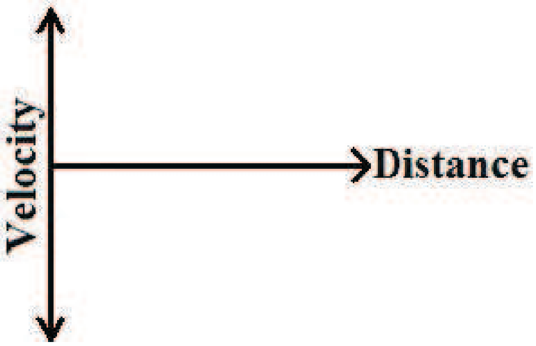


Figure 1 at right, is a possible graph showing how fast galaxies move away from us in the expanding universe.

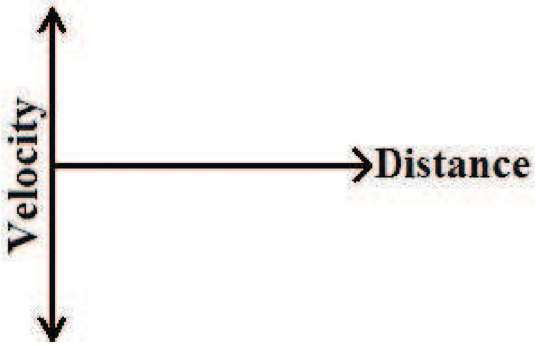
5) Use the blank graph provided below to draw what your think Figure 1 would look like if the universe had been expanding twice as fast. Explain the reasoning behind the graph you drew.

   If you don't have enough information to do this, explain what else you need to know.



6) Use the blank graph provided below to draw what your think Figure 1 would look like for a much older universe. Explain the reasoning behind the graph you drew.

   If you don't have enough information to do this, explain what else you need to know.



7) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

8) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

D.2 Form B

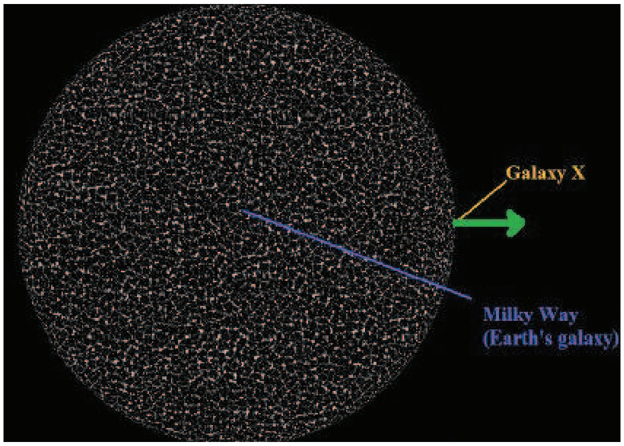
The spring 2010 version of Form B begins on the next page.

Name: ________________________________ **Conceptual Cosmology Survey - Form B**

1) Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." Provide a drawing if possible to help illustrate your thinking.

2) Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory." Provide a drawing if possible to help illustrate your thinking.

3) Each dot in the picture on the left is a galaxy. The Milky Way Galaxy (the one we live in) is at the center of the picture. All of the galaxies inside the circle can be seen from Earth. Any galaxies that exist outside the circle are so far away that their light has not had time to reach Earth. Describe what inhabitants of Galaxy X probably see when they look in the direction of the arrow.



4) Circle the phrase that best completes the sentence. Surrounding the event called the BB was ____.
a) a region of space that includes nothing (empty space)
b) a region of space that includes particles and matter
c) neither a nor b
Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

5) Independent of whether we know its true location, is there a center to the universe? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

6) If you could travel to any location in the universe, could you go to a place where there are no galaxies in front of you? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

7) Which of the following statements (a - d) are true? Circle all that apply.
In general, the expansion of the universe causes the distances between ________.
a) planets in the solar system to increase.
b) stars in the galaxy to increase.
c) galaxies in the universe to increase.
Explain your reasoning for your choice(s).

8) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

9) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

D.3 Form C

The spring 2010 version of Form C begins on the next page.

Name: ___________________________________ **Conceptual Cosmology Survey - Form C**

1) Over time, would you say the temperature of the universe has increase, decreased, or stayed the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

2) How does the total amount of matter (not energy) in the universe *right now* compare to the total amount of matter at the *very beginning* of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

3) Over time, would you say the overall density of matter in the universe has increase, decreased, or stayed the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

Questions 4-6 refer to this situation: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X.

4) How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y?
a) less than 8 billion light-years apart b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

5) How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode?
a) less than 8 billion light-years apart b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

6) The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age?

a) less than 5 billion years old b) exactly 5 billion years old
c) more than 5 billion years old d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

7) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

8) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

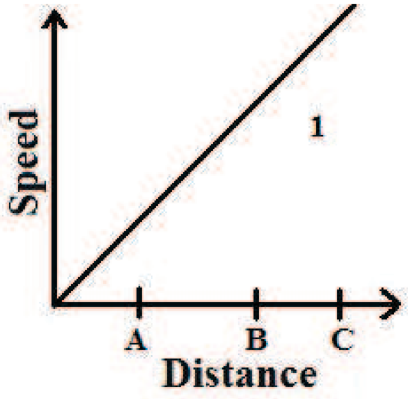
D.4 Form D

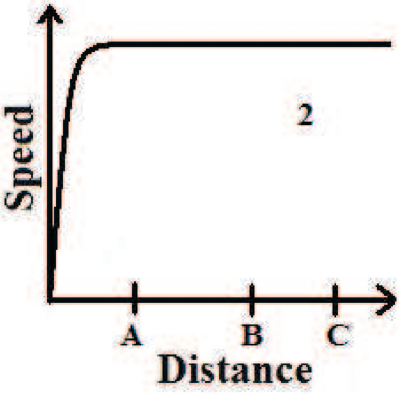
The spring 2010 version of Form D begins on the next page.

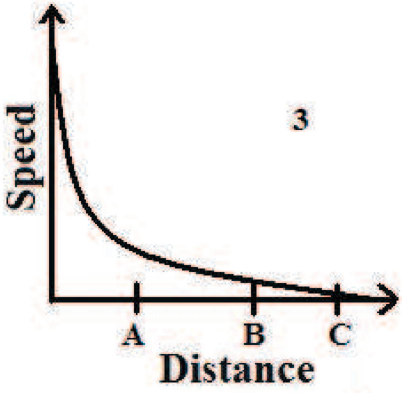
Name: ______________________________ **Conceptual Cosmology Survey - Form D**

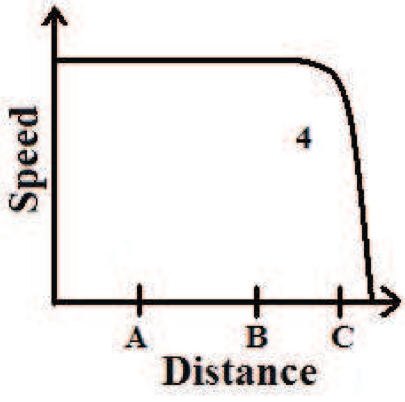
1) Here is a picture of a spiral galaxy. The locations of three stars (A - C) are shown.

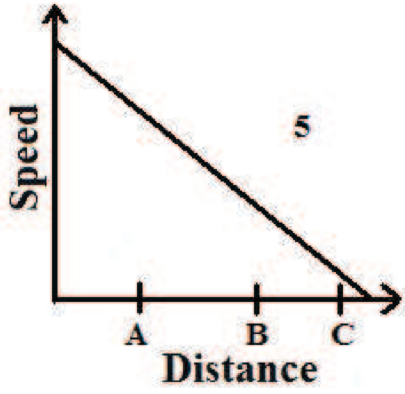
When we observe how fast stars orbit the center of the galaxy versus their distance from the center, we plot the results in what's called a *rotation curve*. Below are some possible rotation curves (1-6). The locations of Stars A-C are shown on the distance axis of each graph.

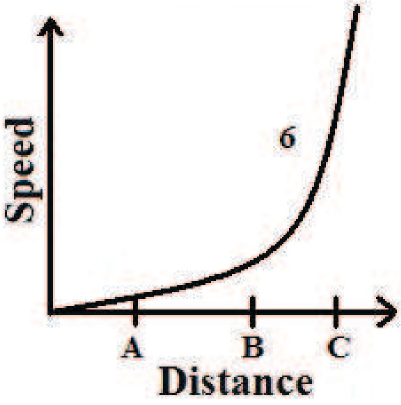












1) Which graph best represents how stars orbit the center of our galaxy? Explain the reasoning behind your choice.

2) Rank the speeds at which Stars A, B, and C orbit the galaxy.

Ranking Order: Fastest speed 1________ 2________ 3________ Slowest speed

Or, all the stars orbit at approximately the same speed. _____ (indicate with a check mark)

Carefully explain your reason for ranking this way:

3) Rank the net gravitational force exerted on Stars A, B, and C.

Ranking Order: Strongest force 1________ 2________ 3________ Weakest force

Or, all the stars feel approximately the same force. _____ (indicate with a check mark)

Carefully explain your reason for ranking this way:

4) Based on your previous answers, how is matter distributed in our galaxy? Pick the best answer from the following choices (a-c).
a) Most of the matter in the galaxy is located in the center.
b) Most of the matter in the galaxy is located in the center and spiral arms.
c) Neither a nor b.
Explain your reasoning for your choice.

5) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

6) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

Appendix E

Spring 2010 Lecture-Tutorials

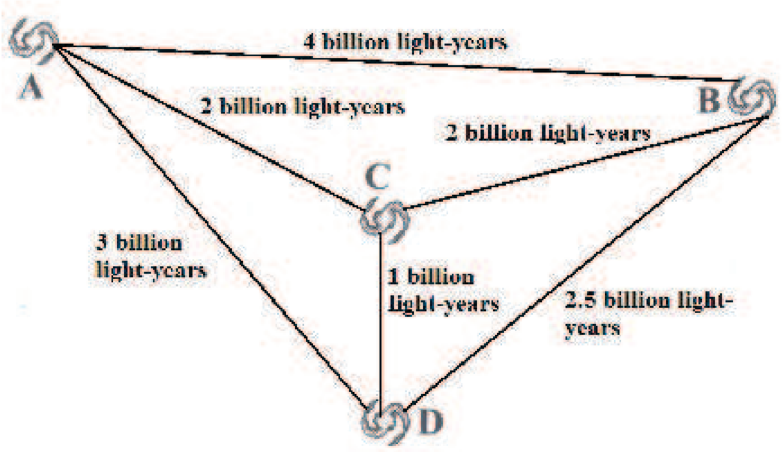
E.1 Hubble's Law

The spring 2010 version of the “Hubble's Law” lecture-tutorial begins on the next page.

Hubble's Law

Consider the small section of the universe containing four galaxies (A-D), shown in Figure 1 below. The distances between each galaxy are also shown.

Figure 1



1) Imagine that this section of the universe doubles in size over time due to the expansion of the universe. Draw what the above section of the universe would look like after it doubles in size. Be sure to identify the new distances between the galaxies.

Hubble's Law

2) Which of the galaxies (B-D) increased its distance from Galaxy A by the greatest amount (greatest number of light-years) during this time? Explain your reasoning.

3) Two students are discussing their answers to Question 2:

Student 1: *All of the distances doubled, so all of the distances increased by the same amount. There is no one galaxy whose distance from Galaxy A increased the most.*

Student 2: *You're right that all the distances double in size, but I don't agree that they all increase by the same number of light-years. Since Galaxy B was the farthest away from Galaxy A initially, its distance will increase by the greatest number of light-years when this section of the universe doubles in size.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

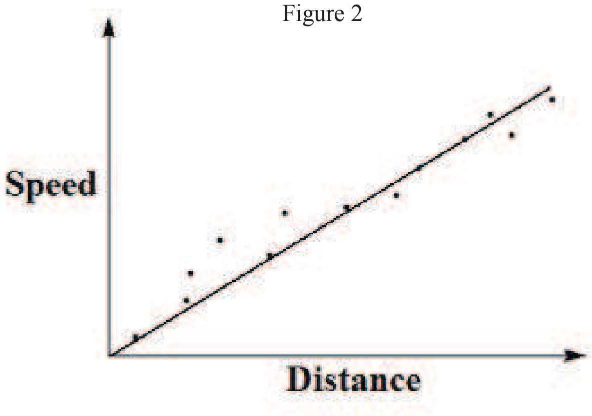
4) Describe the relationship between a galaxy's distance from Galaxy A and the speed at which it appears to be moving away from Galaxy A. Explain your answer in terms of your answers to Questions 1-3.

5) Is the relationship you described in Question 4 unique to Galaxy A, or would you observe the same relationship if you lived in one of the other galaxies? Explain your reasoning.

The relationship you described in Questions 4 and 5 is called *Hubble's law*. We can depict Hubble's law with the graph shown below. This graph plots the speed at which a galaxy appears to move away from us versus its distance from us. This type of graph is called a *Hubble plot*. Each dot on the plot represents a different galaxy.

Hubble's Law

Figure 2



6) Explain how the Hubble plot shown in Figure 2 above is consistent with the relationship you described in Question 4.

7) Imagine the Hubble plot shown in Figure 2 represents a universe that doubles in size over a certain amount of time. Which of the two Hubble plots shown in Figures 3 and 4 below might represent a universe that is tripling in size over the same amount of time? Explain your reasoning.

Figure 3

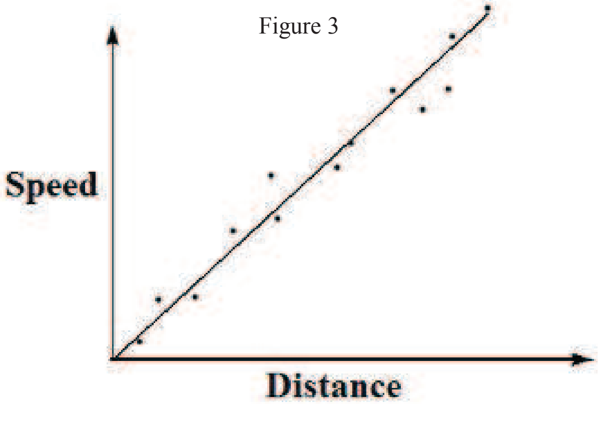
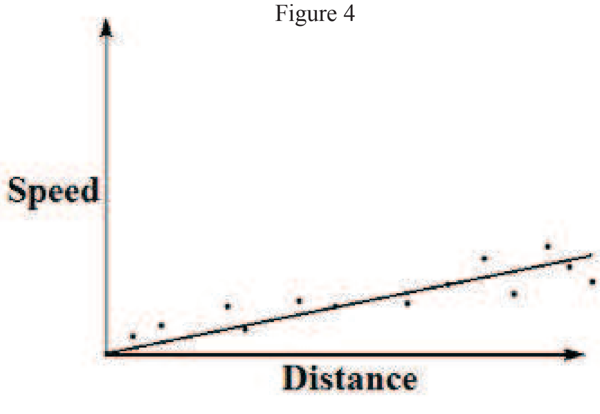


Figure 4



Hubble's Law

The *expansion rate* of the universe determines how fast it is growing in size. For example, a universe that is tripling in size has a faster *expansion rate* than a universe that is doubling in size over the same amount of time. In a Hubble Plot, the *expansion rate* is indicated by the slope of the graph.

8) Would you say the expansion rate for the universe represented in Figure 2 is constant, increasing, or decreasing?

9) Rank the expansion rates (from fastest to slowest) of the three different universes represented in Figures 2, 3, and 4. Explain your reasoning.

10) If the expansion rate of our universe had been faster, would the universe have reached its current size earlier in its history or later? Explain your reasoning.

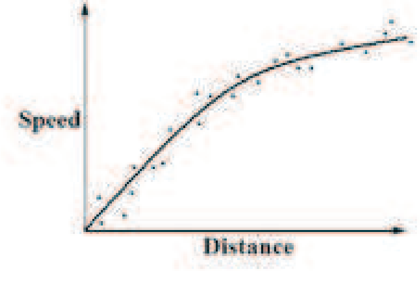
11) Complete the sentence below using the words provided in parentheses ().

For two universes that are the same size, the universe with the faster expansion rate must be __________(younger/older) than the universe with the slower expansion rate.

12) If the above Hubble plots (Figures 2-4) represent three universes that are the same size, which Hubble plot belongs to the youngest universe? Explain your reasoning.

Recent observations indicate the Hubble plot for our universe actually looks more like the plot in Figure 5.

Figure 5



Hubble's Law

13) Based on the straight line drawn in the Hubble plots shown in Figures 2-4, you might infer that the expansion rate for the universe is constant. Based on the Hubble plot shown in Figure 5, would you say that the expansion rate of the universe is constant or changing over time? Explain your reasoning.

14) Based on the Hubble plot in Figure 5, is the expansion rate represented by the motion of galaxies far away from us faster than, slower than, or the same as the expansion rate represented by motion of nearby galaxies? Explain your reasoning.

15) Based on the Hubble plot in Figure 5, is the expansion rate of the universe increasing or decreasing as time goes on? Explain your reasoning.

16) Consider the following debate between two students regarding their answer to the previous question:

Student 1: *The slope of the graph tells you how fast the expansion rate of the universe is, not how fast a galaxy is moving. The farther we look into space, the further we are looking back in time. Since the slope of the Hubble plot is flatter in the past and steeper now, that means the expansion rate has sped up over time.*

Student 2: *I think you are reading the graph wrong. The expansion rate of our universe must be slowing down as time goes on. If you look at the Hubble plot, you can see that the graph gets flatter. That means the farther away you look, the slower the expansion rate is. The rate at which the most distant galaxies are moving away from us has started to slow down and eventually the expansion rate of nearby galaxies will also slow down.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

Hubble's Law

Distant galaxies appear dimmer in our sky than closer galaxies. This is because we receive less light from distant galaxies than we do from nearby galaxies. So the farther away a galaxy is from us, the harder it is to see.

17) Complete the sentence below using the words provided in parentheses ().

    Our observable universe is getting __________ (darker/brighter) over time.

    Explain your reasoning.

18) Has the rate at which the universe is getting darker or brighter been the same for the entire age of the universe or is it getting darker at a faster (or slower) rate now as compared to the past? Explain your reasoning.

19) Is the above answer true for all locations in the universe? Explain your reasoning.

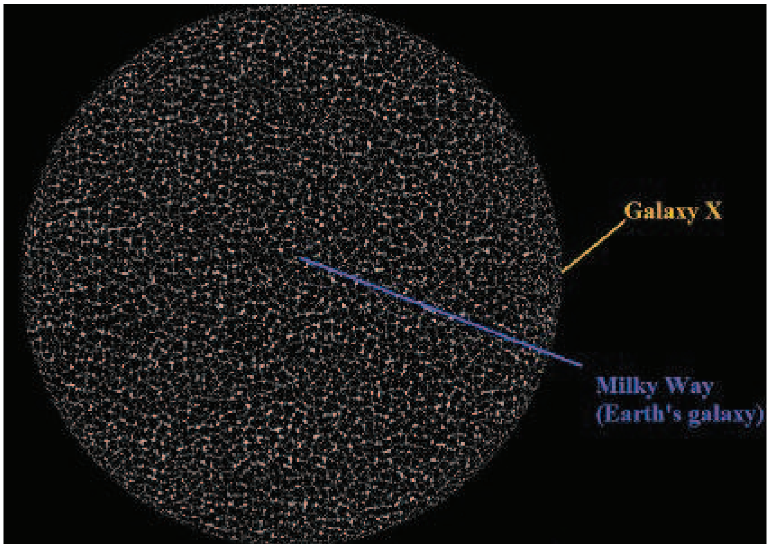
E.2 Making Sense of the Universe and Expansion

The spring 2010 version of the "Making Sense of the Universe and Expansion" lecture-tutorial begins on the next page.

Making Sense of the Universe and Expansion

Part I: The Observable Universe

Each dot in the picture below represents a galaxy. The Milky Way galaxy is represented by the dot at the center of the picture. All of the galaxies inside the circle can be seen from Earth. The circumference of this circle defines what is called our *observable universe*. Any galaxy that exists outside the circle is so far away that its light has not had time to reach Earth and is therefore not part of our observable universe.



1) Do you think the galaxies we can see from Earth are the only galaxies in the universe? Explain your reasoning.

2) Draw a circle around Galaxy X that represents its observable universe.

3) Is the observable universe that you drew for Galaxy X different in size than the observable universe for the Earth? Explain your reasoning.

Making Sense of the Universe and Expansion

4) Two students are talking about the observable universe for Galaxy X:

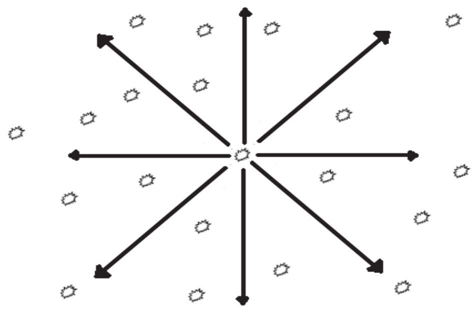
Student 1: *People living in Galaxy X have a strange view of the universe. When they look in one direction they see a bunch of galaxies, but when they look in the other direction all they see is empty space. Galaxy X must be at the edge of the universe since there's nothing but black, empty space beyond it. We're lucky we live at the center since we can see galaxies all the way out to the edge of the universe, no matter where we look.*

Student 2: *I think you're wrong. People living in Galaxy X would probably see a bunch of galaxies in every direction they look, but they can see some galaxies that we can't, just like we can see galaxies they can't. The observable universe for any galaxy should look similar to ours. I don't think we are at the center of the universe and I don't think Galaxy X is at the edge either.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

Part II: An Analogy for Expansion

Consider the following description and sketch from a student regarding the expansion of the universe.

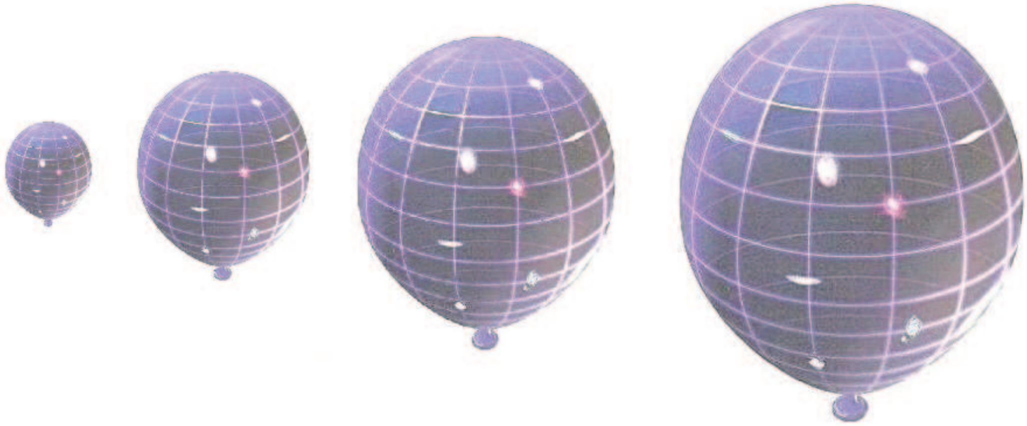


Student: *The expansion of the universe means that all galaxies are moving away from the center of the universe. If you lived at the center, you'd see all the galaxies moving away from you. Since this is what we see, we must be at the center. Some galaxies are at the edge of the expansion. They're moving into empty space, so if you lived in one of the galaxies at the edge of the universe, you'd see that you are moving into a region where there are no galaxies.*

Making Sense of the Universe and Expansion

5) Do you agree or disagree with this student's description about the center and edge of the expanding universe? Explain your reasoning.

One way to try to understand and envision the expansion of the Universe is by creating analogies that attempt to model the different aspects of our real expanding Universe. One way to model the expanding Universe is to use a "balloon" analogy. In this analogy, the space and time of the universe are modeled by the "surface" or "skin" of an expanding balloon. Galaxies only exist on the surface of the balloon. Light can only travel on the surface of the balloon.



6) Do objects, light, or events in the universe also exist inside or outside of the balloon's surface in this analogy?

7) If you were to travel from galaxy to galaxy along the surface of the balloon, would you ever encounter an edge? Explain your reasoning.

8) If you were to travel across the entire surface of the balloon universe, would you ever find a location that you would consider to be the center of the universe? Explain your reasoning.

Making Sense of the Universe and Expansion

9) Consider the following debate between two students about their answers to the previous questions:

Student 1: *Someone living on the surface of this balloon universe will definitely encounter an edge and a center. All they have to do is look from their location across the inside of the balloon to a location on the other side. The center of the inside of the balloon is the center of the universe, and the far side would be the edge of what they could see. So there's definitely a center and an edge to the universe in the balloon analogy.*

Student 2: *I think you misunderstand the analogy. The surface of the balloon is supposed to be the entire universe. The inside of the balloon isn't part of the universe. You can't look through the inside of the balloon to the other side so there is no center in the middle or edge on the other. In this analogy, people living in the balloon universe would never see a center or an edge.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

10) Imagine you lived in a galaxy on the surface of the balloon. As the balloon expands, would all the other galaxies appear to move toward you or away from you?

11) Would your answers to the previous question be the same regardless of the galaxy in which you live, or would it change depending on the galaxy you inhabit?

12) In this analogy, do galaxies move relative to one another because they are traveling across the surface of the balloon, or do they move relative to one another because the balloon is expanding?

Making Sense of the Universe and Expansion

13) The balloon analogy is a helpful way to think about expansion, but no analogy is perfect. Some aspects of the real universe are captured by this analogy while others are not. Below are several properties of the real universe. For each, state whether it is accurately captured by the balloon analogy or not, and explain your reasoning.

a) The real universe has no center.

b) The real universe has no edge.

c) The real universe is expanding.

d) Light can travel in straight lines in the real universe.

e) The real universe's expansion does not cause galaxies to change size.

f) The real universe is 4-dimensional (3 space dimensions and 1 time dimension).

E.3 Expansion, Lookback Times, and Distances

The spring 2010 version of the "Expansion, Lookback Times, and Distances" lecture-tutorial begins on the next page.

Expansion, Lookback Times, and Distances

When the universe was 4 billion years old, Galaxy A was 3 billion light-years away from Galaxy B, as shown below. Imagine that the universe was not expanding, so the distance between Galaxy A and Galaxy B would not change over time.

1) A star explodes in Galaxy B producing a large amount of light. How long will the light from this explosion take to reach Galaxy A?

2) How far did the light travel on its journey to Galaxy A?

3) How old will the universe be by the time the light from the explosion reaches Galaxy A?

Because light takes time to travel from place to place in the universe, when we look at the night sky we are seeing stars and galaxies as they appeared in the past. For example, if we see a galaxy 1 million light-years away, we are seeing what the galaxy looked like 1 million years ago. We often use the term *lookback time* to describe how far back in time we are seeing. So if we see a galaxy as it appeared 1 million years ago, we say it has a lookback time of 1 million years.

4) What is the lookback time inhabitants of Galaxy A associate with Galaxy B when they see the light from the explosion?

The real universe is expanding. This means the distance between galaxies is constantly increasing. Imagine that Galaxy A and Galaxy B are in an expanding universe.

5) While the light from the explosion is traveling from Galaxy B to Galaxy A, does the distance between the two galaxies stay the same, become larger, or become smaller?

Expansion, Lookback Times, and Distances

6) By the time the light from the explosion reaches Galaxy A, is the distance to Galaxy B more than, less than, or exactly 3 billion light-years?

7) By the time the light from the explosion reaches Galaxy A, has more than, less than, or exactly 3 billion years elapsed since the star exploded?

8) By the time the light from the explosion reaches Galaxy A, will the total distance traveled by the light be more than, less than, or exactly 3 billion light-years?

9) When the inhabitants of Galaxy A see the light from the explosion, are they looking at an event with a lookback time of more than, less than, or exactly 3 billion years in the past?

10) In the space below provide a sketch that explains the reasoning behind your answers to questions (6-9).

Expansion, Lookback Times, and Distances

11) Consider the discussion between two students regarding their ideas about two distant galaxies in an expanding universe.

Student 1: *Let's say light takes 5 billion years to travel from one galaxy to another. This means the two galaxies were separated by 5 billion light-years when the light began its journey.*

Student 2: *If the light traveled for 5 billion years, then the distance between the two galaxies must have been less than 5 billion light-years when the light began its journey because the distances between galaxies are always increasing in the expanding universe.*

Do you agree with either or both of the students? Explain your reasoning.

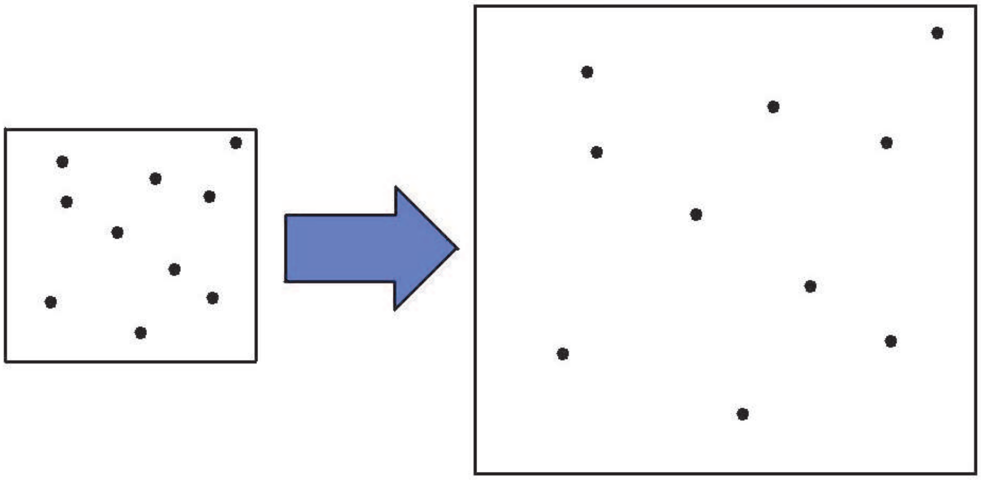
E.4 The Big Bang

The spring 2010 version of “The Big Bang” lecture-tutorial begins on the next page.

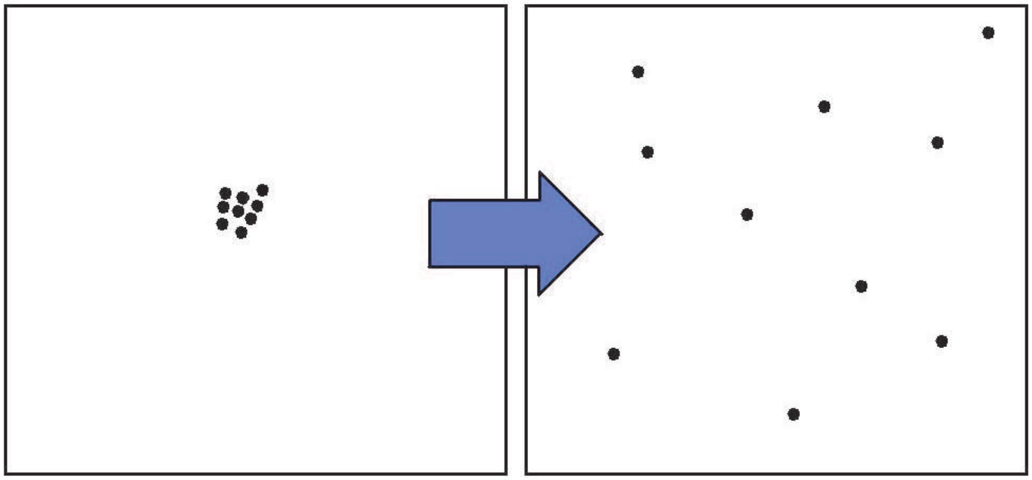
The Big Bang

Consider the drawings (A and B) provided below which each represent a different way of thinking about how the universe changes over time. The dots in each diagram drawing represent pieces of matter.

A



B



1) Which drawing, A or B is a better representation of the universe we observe? Explain your reasoning.

The Big Bang

2) In Diagram A, is the universe becoming bigger, smaller, or staying the same size over time?

3) In Diagram B, is the universe becoming bigger, smaller, or staying the same size over time?

4) Two students are debating their answers to Questions 2 and 3:

Student 1: *Both diagrams show the universe becoming bigger. In Diagram A, the box has become larger. In Diagram B, the pieces of matter have spread out.*

Student 2: *I disagree. Only Diagram A shows the universe becoming bigger. In Diagram B the size of the box doesn't change. The pieces of matter are just moving into an already existing empty space in a universe whose size doesn't change.*

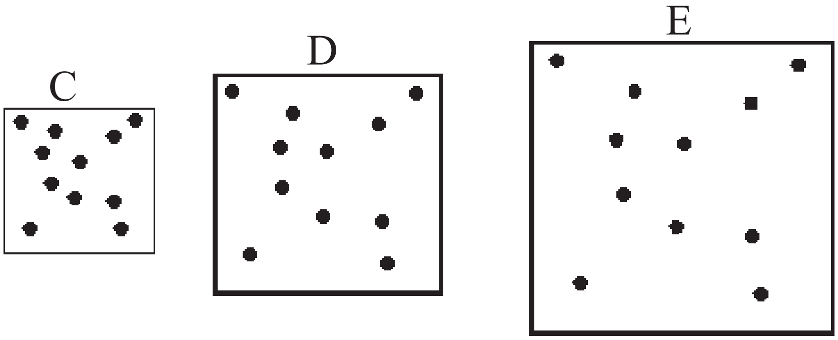
Do you agree or disagree with either or both of the students? Explain your reasoning.

5) Both drawings show the distance between matter increasing over time.
a) Which of the drawings shows this happening as the result of space expanding and which is a result of an outward explosion?

b) Which of the drawings would you say provides a more correct representation of our universe? Is your answer to this question consistent to your answer to question #1? Explain your reasoning.

The Big Bang

Consider the three drawings (C, D and E) shown below. These drawings each represent the same region of space, but at different times during the history of the universe.



6) Which drawing shows the region of space at the earliest point in the history of the universe? Explain your reasoning.

7) In which drawing does the region of space have:
a) the highest density?

b) the greatest concentration of energy?

c) the highest temperature?

Explain your reasoning.

8) Imagine you could watch the history of the universe like a movie playing backward. The movie starts today and ends at the beginning of the universe. Describe what you would see as the movie played and you looked further back in time. Your answer should discuss how the universe changes in terms of its temperature, density, and size.

The Big Bang

Your answers to the previous questions are all part of the *Big Bang Theory*. The Big Bang Theory does not say what the universe was like at the very first moment of time, which was about 13.7 billion years ago. It does, however, tell us how the universe changed after its first moment of existence.

9) Three students are discussing their understandings of the Big Bang Theory:

Student 1: *I think I understand the Big Bang now. At the beginning, all the matter in the universe was compacted into a small, hot, dense ball. This ball of matter then exploded into empty space. When we look at the universe, we see galaxies moving away from us. The Big Bang model explains this, since all matter should be flying away from the center point of the explosion.*

Student 2: *I disagree. I think what the Big Bang Theory is saying is that all the matter in the universe was once compacted into a really small, dense and hot object that expanded over time. But, there wasn't an explosion of matter into empty space. Instead, the universe carried galaxies and other matter away from each other as it expanded in size.*

Student 3: *You're both wrong. I agree that the universe was once smaller in size and that pieces of matter have been carried away from each other by the expansion of the universe. But remember how we learned from Einstein's equation* $E = mc^2$ *that matter can be converted into energy and energy can be converted into matter? I think this means that if we go back to the beginning of the universe, it would be so small, and its temperature so hot that matter itself can't exist. I bet at the very beginning, the universe would have been infinitely small and composed of pure energy with no matter there at all.*

Which students do you agree or disagree with? Explain your reasoning.

The Big Bang

10) Based on your previous answers, complete the following sentences:

    The Big Bang Theory says that the universe started out with a very __________ temperature, a very ___________ density, and a very ___________ size. Originally, there was ______ matter, only pure ___________. From this initial state, the universe ____________ in size. This caused its temperature and density to ___________. When the temperature was cool enough, energy could transform into ___________.

11) Look at drawing A again. Next to drawing A, make a drawing of what you think the universe would have looked like at the very first instant it existed.

E.5 Dark Matter

The spring 2010 version of the “Dark Matter” lecture-tutorial begins on the next page.

Dark Matter

Part I: Motions of Planets

An object's orbit depends on the "mass inside" its orbit (also known as *interior mass*). For a planet in our Solar System, you can find the interior mass by adding the Sun's mass to the mass of each object between the Sun and the planet's orbit. For example, the interior mass to Earth's orbit would be the Sun's mass plus the mass of Mercury plus the mass of Venus.

Here is a table that lists each planet, the mass inside each planet's orbit, and the speed at which the planets orbit the Sun.

Planet	Interior Mass (solar masses)	Orbital Speed (km/s)
Mercury	1.00	47.9
Venus	1.0000027	35.0
Earth	1.0000057	29.8
Mars	1.0000060	24.1
Jupiter	1.00096	13.1
Saturn	1.0012	9.66
Uranus	1.0013	6.81
Neptune	1.0013	5.43

1) Where is the vast majority of mass in the solar system located? What object or objects account for most of this mass?

2) How does the orbital speed of planets farther from the Sun compare to the orbital speed of planets closer to the Sun?

The mass inside a planet's orbit affects how fast the planet moves because it affects the strength of the gravitational force felt by the planet. In addition, the size of the planet's orbit can affect the strength of the gravitational force exerted on the orbiting planet and therefore affects the planet's orbital speed.

3) How does the gravitational force on planets farther from the Sun compare to the gravitational force on planets closer to the Sun? Explain your reasoning.

Dark Matter

4) Complete the blanks in the sentences of the following paragraph by either writing in the necessary information or circling the correct response. It may be helpful to base your responses on the information provided in the table above and your answers to the previous questions.

   There are ___ planets inside Neptune's orbit and ___ planets inside Mercury's orbit. However, the interior mass for Neptune is ___ (much greater than/approximately the same as/much less than) *the interior mass of Mercury. Neptune is* ___ (much closer to/much farther from/about the same distance from) *the Sun as/than Mercury. Therefore the gravitational force exerted on Neptune is* ___ (stronger/weaker/about the same strength) *as/than the force exerted on Mercury. As a result, Neptune has an orbital speed that is* ___ (much slower, much faster, about the same speed) *as/than the orbital speed of Mercury.*

If you could increase the amount of mass inside a planet's orbit, you would increase the gravitational force it would feel and thus increase the orbital speed of the planet. Imagine you were able to add a very, very large amount of mass distributed evenly *between* the orbits of Jupiter and Saturn.

5) Which planet(s) will experience an increase in gravitational force and an increase in orbital speed from this added mass? Explain your reasoning.

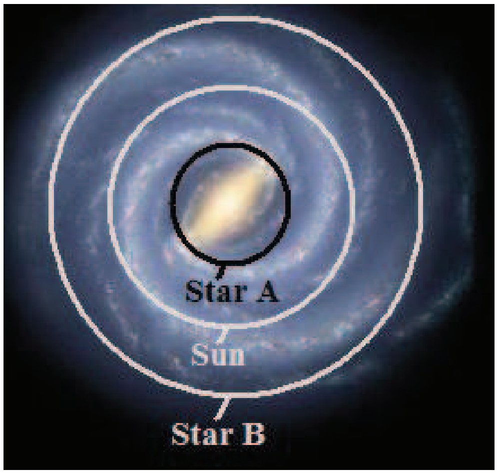
Part II: Motions of Stars

One way to estimate the amount of mass in a spiral galaxy is by looking at how much light the galaxy gives off. Where there is more light there must be more stars and hence more mass. When astronomers measure the amount of light coming from different parts of the galaxy, they find that stars are very concentrated in the central bulge of the galaxy and more spread out as you look outward through the disk.

6) Based on the information provided above, where do you expect most the mass of a galaxy to be concentrated?

Dark Matter

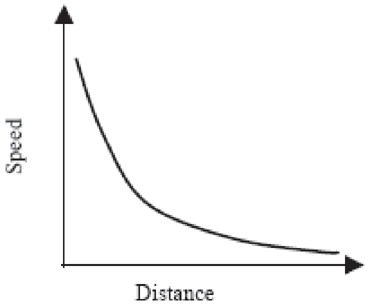
At right is a drawing of the Milky Way. The orbits of three stars are labeled. Star A is a star on the edge of the Milky Way's bulge. The Sun's orbit is shown at approximately the correct position. Star B is a star located farther out in the disk than the Sun.



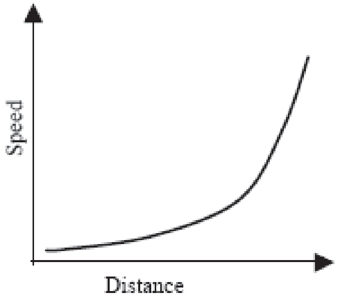
7) Based on your answer to question 6 and the location of each star from the center of the galaxy, rank how you think the force of gravity on Star A, Star B, and the Sun would compare from strongest to weakest. Explain your reasoning.

8) Based on your answer to questions 6 and 7, rank how you think the orbital speeds of Star A, Star B, and the Sun would compare from fastest to slowest. Explain your reasoning.

A graph of the orbital speed of stars versus their distance from the galaxy's center is called a *rotation curve*. Here are two possible rotation curves.



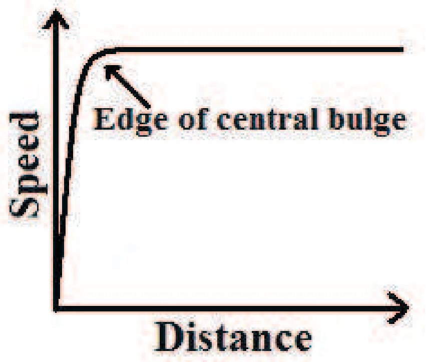
Curve 1



Curve 2

9) Which of the above rotation curves best represents the relationship you described in Question 8? Explain your reasoning.

Dark Matter



Astronomers have measured how fast stars actually orbit around the center of our galaxy. At right is the rotation curve for the Milky Way Galaxy (MWG).

10) Based on their orbital distance from the center of the galaxy, make three dots on the above rotation curve to represent Star A, Star B, and the Sun. Be sure to label which mark belongs to each star.

11) Describe how the real rotation curve for the MWG is different from the rotation curve that you chose in Question 9.

12) Using the real rotation curve for the MWG, provide a new ranking for the orbital speed of Star A, Star B, and the Sun, from fastest to slowest. Describe any differences between this ranking and the one you provided in question 8.

13) Based on your answer to questions 11 and 12, provide a new ranking for the gravitational force on Star A, Star B, and the Sun, from strongest to weakest. Explain your reasoning.

14) Based on your answers to question 10-13, would you say that the mass of the Milky Way Galaxy is concentrated at its center (as is the case with our Solar System)? Explain your reasoning.

15) Two students are debating their answers to the previous questions:

Student 1: *Stars far from the center of the Milky Way are all moving at about the same speed. This means there must be more mass throughout the outer regions of the galaxy than we can see. If there were not, Star B would be moving slower than the Sun.*

Dark Matter

Student 2: *I disagree. There are fewer stars in the outskirts of the Milky Way than in the center, so there's less mass out there than at the center. Most of the Milky Way's mass must be at its center. So I think the location of the mass doesn't affect how fast stars are moving.*

Do you agree with either or both of the students? Explain your reasoning.

16) Astronomers initially thought there was more mass in the center of the galaxy than in the disk because there's more light coming from the galaxy's center than its disk. What observations have astronomers made that show these initial ideas about how mass is distributed in the galaxy are wrong? Explain your reasoning.

17) Is there more or less mass in the Milky Way's disk and halo than we can see? Explain your reasoning.

Appendix F

Spring 2010 Scoring Rubrics

The scoring rubrics for the spring 2010 begin on the next page.

Form A

Item 1: Which graph or graphs show a universe that is expanding at a constant rate? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the expansion (or contraction) rate of the universe.
>
> **Correct answer:** F - This graph shows an expanding universe because galaxies that are farther away are moving away from us (positive velocity) faster than closer galaxies. Since the slope is constant, the rate at which the universe is expanding must be constant.

Item 2: Which graph or graphs show a universe that is contracting at a constant rate? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the expansion (or contraction) rate of the universe.
>
> **Correct answer:** B - This graph shows a contracting universe because galaxies that are farther away are moving toward us (negative velocity) faster than closer galaxies. Since the slope is constant, the rate at which the universe is contracting must be constant.

Item 3: Which graph or graphs show a universe that is expanding at a faster and faster rate over time? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the changing expansion (or contraction) rate of the universe, relative to lookback time.
>
> **Correct answer:** A - The rate at which the universe is expanding must be changing over time since the slope of the graph is changing with respect to distance. A universe expanding faster and faster over time must have a slope that gets steeper over time. This means the slope must be flatter at large distances since we are looking further back in time as we look farther away in distance.

Item 4: Which graph or graphs show a universe that is expanding at a slower and slower rate over time? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the changing expansion (or contraction) rate of the universe, relative to lookback time.
>
> **Correct answer:** D - The rate at which the universe is expanding must be changing over time since the slope of the graph is changing with respect to distance. A universe expanding slower and slower over time must have a slope that gets flatter over time. This means the slope must be steeper at large distances since we are looking further back in time as we look farther away in distance.

The scoring rubric for these four items is on the next page.

Scoring rubric for items 1a-d.

Overall Score	Answer Choice	Reasoning Categories
3 - Student gives the correct answer with the correct reasoning.	A - Student selects graph A.	I - Description includes all or part of the following statement: "†̲ is ±̲ and is *̲ at a ‡̲." *See below for a complete explanation of this reasoning category and symbols.*
	B - Student selects graph B.	
2 - Student gives the correct answer with incorrect or incomplete reasoning.	C - Student selects graph C.	II - Student talks about how when we look deep into space we are looking back in time.
	D - Student selects graph D.	III - Student says her/his reasoning is opposite his/her reasoning on the previous item.
1 - Student gives an incorrect answer.	E - Student selects graph E.	IV - Student says s/he is guessing and/or unsure of why s/he chose that answer.
	F - Student selects graph F.	
0 - Student gives no answer, offers irrelevant information, or claims s/he has no idea how to answer.	G - Student selects graph G.	V - Student says s/he can't give an answer without more information.
	H - Student selects graph H.	VI - Student gives no reasons for answer.
	Z - Student selects no graph.	VII - Student gives some other reason not captured by categories I-X above.

Each response receives an overall score (0-3) as when as a letter (A-H or Z) to denote which graph the student chose. A Roman numeral (I-VII) is also assigned to each response to mark the student's reasoning. Reasoning Category I is further subdivided, as described on the next page.

Reasoning Category I explained - Many students justify their selections by noting that the graph or some aspect of the graph is 1) positive or negative (i.e. above or below the x-axis), 2) increasing, decreasing, constant, or curved, and 3) changing with an increasing, decreasing, or constant rate. Some students include all of these elements in the reasons they wrote; most only have a subset of these reasons. To better distinguish between these various possibilities, there are various subclasses for Reasoning Category I:

When the student writes "† is ± and is * at a ‡," append, when appropriate, the following symbols to I.

† = v *if the student discusses velocity or speed*

= d *if the student discusses distance*

= s *if the student discusses slope*

= r *if the student discusses rate*

= ℓ *if the student discusses the line*

= g *if the student discusses the graph*

= o *if the student discusses some other feature*

= i *if the student discusses an unspecified feature or "it"*

± = + *if the subject † is positive (above the x-axis)*

= - *if the subject † is negative (below the x-axis)*

* = INC *if the subject † is increasing*

= DEC *if the subject † is decreasing*

= CON *if the subject † is constant/straight*

= EXP *if the subject † is curved or exponentiating*

‡ = R1 *if the rate at which subject † changes increases*

= R2 *if the rate at which subject † changes decreases*

= R3 *if the rate at which subject † changes stays constant*

Example responses and scores:

A response to item 1a: *"F. The line is straight, making the expansion constant and it is increasing both Velocity and Distance."*

> **How this response is scored:** 3F_I_lCON_vINC_dINC. This student gave the correct answer (F) and a correct reason. She thus receives an overall score of 3. In her answer, she discusses how the line (l) is constant (CON), the velocity (v) is increasing (INC), and the distance (d) is increasing (INC).

A response to item 1b: *"B, graph constantly going down."*

How this response is scored: 2B_I_gDEC_R3. This student gives the correct answer (B) but with an incomplete reason, thus leading to an overall score of 2. In her answer, she says the graph (g) is going down (DEC) constantly (R3).

A response to item 1c: *"D"*

How this response is scored: 1D_VI. This student gave the wrong answer (D, hence leading to an overall score of 1) and gave no reason (VI).

A response to item 1d: *"A,B,E. Velocity is decreasing as the distance is increasing"*

How this response is scored: 1_ABE_vDEC_dINC. This person earns an overall score of 1 because he gave the wrong answer (A, B, and E). His reason talks about the velocity (v) decreasing (DEC) and the distance (d) is increasing (INC).

Item 5: Use the blank graph provided below to draw what your think Figure 1 would look like if the universe had been expanding twice as fast **.** Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

What this item measures: This item probes whether or not a student knows that the slope of a line on a Hubble plot should steepen if the expansion rate increases.

Scoring rubric for item 5.

Score	Response Includes	Reasoning Elements
2	- correct answer (steeper slope).	I - Student draws or discusses a steeper slope.
1	- incorrect answer or incorrect or incomplete reasons for choosing the correct answer.	II - Student draws or discusses a flatter slope. III - Student draws or discusses an unchanged slope. IV - Student draws a line with a non-constant slope.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	V - Student draws a line whose slope is ≤ 0. VI - Student draws a line whose y-intercept ≠ 0. VII - Student talks about increased velocity. VIII - Student talks about decreased velocity. IX - Student talks about velocity staying the same. X - Student talks about increased distance. XI - Student talks about decreased distance. XII - Student talks about distance staying the same. XIII - Student compares variable *x* to variable *y* by saying *x* is more than *y.* XIV - Student says distance and velocity change by the same amount. XV - Student says s/he can't answer the question without a time axis/variable. XVI - Student gives irrelevant information. XVII - Student gives some other reason not specified above. XVIII - Student has no idea. XIX - Answer field is blank.

Example responses and scores:

"The line would be much steeper b/c for every 1 unit of distance it expanded, the velocity would increase twice as much." [Student also draws a straight line with a steeper slope than Figure 1.]

> **How this response is scored:** 2_I_VII_XIII_X. This student draws a steeper line and says the line should be steeper (I). She also says both the distance (X) and velocity (VII) increase, although velocity increases more than distance (XIII).

"Same line, but units on graph will be twice as much" [Student also draws a straight line with the same slope as Figure 1.]

> **How this response is scored:** 1_III_XIV. This student drew and discussed a slope that was the same as in Figure 1 (III). He also said that the values of both velocity and distance double

(XIV).

Item 6: Use the blank graph provided below to draw what your think Figure 1 would look like for a much older universe. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

> **What this item measures:** This item probes whether a students knows that, in order for the universe to be older, it would need longer to reach its current size, which in turn means the expansion rate and slope of the Hubble plot both have to be smaller.

Scoring rubric for item 6.

Score	Response Includes	Reasoning Elements
3	- correct answer (flatter slope) - reason includes all of the following: 1) the universe took longer to reach its current size, 2) which means it had a slower expansion rate/velocity, and so 3) the graph has a flatter slope	I - Student draws or discusses a positive slope.
		II - Student draws or discusses a negative slope.
		III - Student draws or discusses a slope of zero.
		IV - Student draws or discusses a non-constant slope.
		V - Student draws or discusses a steeper slope.
		VI - Student draws or discusses a flatter slope.
		VII - Student draws or discusses an unchanged slope.
		VIII - Student draws a line whose y-intercept ≠ 0.
2	- correct answer (flatter slope) - incorrect or incomplete reason	IX - Student talks about a faster expansion rate.
		X - Student talks about a slower expansion rate.
		XI - Student talks about an unchanged expansion rate (i.e. expansion rate is the same as Figure 1).
1	- incorrect answer	XII - Student talks about an expansion rate that speeds up.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	XIII - Student talks about an expansion rate that slows down.
		XIV - Student talks about a constant expansion rate (i.e. expansion rate doesn't change over time).
		XV - Student talks about an increased velocity/speed.
		XVI - Student talks about a decreased velocity/speed.
		XVII - Student talks about a constant velocity/speed.
		XVIII - Student talks about an increased distance.
		XIX - Student talks about a decreased distance.
		XX - Student talks about a constant distance.
		XXI - Student talks about what happened to the universe in the past.
		XXII - Student talks about what will happen to the universe in the future.
		XXIII - Student says there's not enough information.
		XXIV - Students says the universe's age is irrelevant.
		XXV - Student talks about the time the universe needs to reach its current size.
		XXVI - Student says s/he needs to know how the expansion rate changes.
		XXVII - Student says there's no time variable/axis.
		XXVIII - Student says the significance of the universe's age is unclear.
		XXIX- Student gives irrelevant information.
		XXX - Student gives some other reason not specified above.
		XXXI - Student has no idea.
		XXXII - Answer field is blank.

Example responses and scores:

"An older universe would take longer to get to the distance it is today. This means a slower velocity." [Student also draws a straight line with a shallower slope than Figure 1.]

> **How this response is scored:** 3_I_VI_XXV_XVI. This student draws a line with a positive slope (I). The slope is also shallower than Figure 1 (VI). The student talks about the time the universe needs to reach its current size (XXV) and connects this to a slower velocity (XVI).

"A much older universe would have a slower velocity of expansion b/c it may be nearing the time it would start contracting." [Student also draws a straight line with a shallower slope than Figure 1.]

> **How this response is scored:** 2_I_VI_X_XXX. This student draws a line with a shallower slope than Figure 1 (VI). The line's slope is positive (I). She also mentions a slower expansion (X), but then talks about contraction, which is not included in the list of reasoning elements (XXX).

"As the universe gets older, the slope of the line decreases" [Student also draws a graph like Graph A in items 1-4.]

> **How this response is scored:** 1_I_IV_XXII. The student drew a line in which the slope is always positive (I) but is not constant (IV). The student's reasoning describes the future evolution of the universe (XXII).

Form B

Item 1: Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." Provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes students' conceptions of how the universe increases in size. Does a student have an expert-like conception, in which case s/he talks about how galaxies are getting farther from each other due to the stretching (expansion) of space (or spacetime)? Does the student incorrectly view the expansion as affecting the distances between all objects, including stars, planets, and other objects smaller than galaxies? Does the student incorrectly think that the "expansion of the universe" refers to growth in our knowledge or the creation of new objects/locations over time?

The scoring rubric for item 1 is on the next page.

Scoring rubric for item 1.

Score	Response Characteristics	Reasoning Elements
4	Students in category 4 meet the same criteria as students in category 3 EXCEPT they only discuss increasing distances and/or redshifts in the context of galaxies (and not planets, stars, or other objects smaller than galaxies). Students may be placed in category 4 without explicitly saying that • the universe has no center, • the universe has no edge, • the Big Bang refers to the evolution of the universe from a hot, dense state. If the student does discuss one of these elements and makes an incorrect statement then s/he cannot be placed in category 4.	I - Student says the size of the universe increases over time. II - Student says the universe has a center. III - Student says the universe has no center. IV - Student says the universe has an edge. V - Student says the universe has no edge. VI - Student talks about redshifts/Doppler shifts. VII - Student says space(time) is growing/stretching. VIII - Student talks about the movement of galaxies and/or their increasing distances.
3	Students in this category say that • The universe increases in size and/or • The distances between all objects increase. Additionally, they must make at least one of the following claims and not contradict the others: • The universe has no center. • The universe has no edge. • The Big Bang refers to the evolution of the universe from a hot, dense state.	IX - Student talks about the movement of stars and/or their increasing distances. X - Student talks about the movement of planets and/or their increasing distances. XI - Student talks about the movement of objects (something unspecified or not a star, planet, or galaxy) and/or their increasing distances. XII - Student says the distances between everything increase.
2	Students in this category say that the universe increases in size and either provide no other information or claim one or more of the following: • The universe has a center. • The universe has an edge. • The Big Bang is an explosion. • The distances between all objects or objects smaller than galaxies increases.	XIII - Student says farther objects move away faster. XIV - Student talks about the Big Bang. XV - Student talks about an explosion. XVI - Student says the early universe was once hot, small, and/or dense. XVII - Student says we learn more about the universe over time.
1	Students in this category describe the expansion of the universe as something other than the universe getting larger over time. The two most popular answers that fit in this category are • Expansion refers to learning more about the universe over time. • Expansion refers to the creation of new objects over time.	XVIII - Student talks about how we are looking further back in time as we look farther into space. XIX - Student says new things are created in the universe over time. XX - Student gives irrelevant information. XXI- Student gives some other reason not specified above. XXII - Student has no idea.
0	Students in this category write nothing (the answer field is blank), or they provide information that doesn't answer the question, or they say they have no idea and provide no further information.	XXIII - Answer field is blank or the student provided no reason or explanation.

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response will contain every element in a given score category, you must also list the reasoning elements the student uses.

Example responses and scores:

"There was a man named 'hubble' who had a telescope and observed galaxies 'red shifting' or moving away. Now when he observed this he had the best telescope and plotted the rate of speed of expansion. He noticed that although the closer galaxies are red shifting at a higher rate, the most far away galaxies were moving away (Red shifting) at a much higher speed. From this observation astronomers deducted the universe is expanding from all far away galaxies are red shifting At a higher speed"

> **How this response is scored:** 4_VI_VIII_XIII. This student talked about redshifts (VI), how distances between galaxies are increasing (VIII), and how the farther away galaxies are, the faster they move away from us (XIII). Since the student talked about distances to galaxies increasing and did not make any incorrect statements about expansion, he receives an overall score of 4.

"'the universe is expanding' because as time goes on, matter is moving further and further away from other matter. Temperature and density have decreased over time as the universe is expanding and matter has moved away"

> **How this response is scored:** 3_XI_XVI. This student talks about otherwise unspecified pieces of matter moving away from one another (XI) and talks about how the universe used to be hotter and dense (XVI). While none of these pieces of information are necessarily wrong, she only gets an overall score of 3 because she does not specify whether expansion only affects the distances between galaxies or whether the distances between smaller objects are also affected.

"There was a big bang which exploded out everything in the universe. The leading edge of this bang has been expanding ever since."

> **How this response is scored:** 2_XIV_XV_XII_IV. This student talks about the Big Bang (XIV) as an explosion (XV). His answer indicates that distances between everything are increasing (XII) and he mentions an edge (IV). These reasoning elements give him an overall score of 2.

"I don't think that it is acculuy expanding in a physical sense, but instead or knowledge of the univers and the areas that we have discovered is expanding with an increase in technology and invesments in sciens"

> **How this response is scored:** 1_XVII. This student denies that the universe is physically growing. Instead, he says that expansion refers to our increase in knowledge over time (XVII), placing him in category 1.

Item 2: Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory." Provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether students think of the Big Bang as the beginning of the expansion of the universe, NOT an explosion of pre-existing matter into empty space, nor as the origin of something smaller than the universe (the Earth, Solar System, Galaxy, etc.).

The scoring rubric for item 2 is on the next page.

Scoring rubric for item 2.

Score	Response Characteristics	Reasoning Elements
3	Students in category 3 talk about the Big Bang using at least one of the following elements: • It's the beginning of the universe, space, and/or time. • It's when (some) of the elements were created. • It refers to how the universe was once hot, small, and dense. Additionally, they also include at least one of these elements (and do not contradict any of the others): • It was not an explosion. • It was the beginning of expansion. • Matter formed from energy. • It encompassed all of the universe.	I - Student says the Big Bang is the beginning of the universe/everything in universe. II - Student says the Big Bang is the beginning of expansion. III - Student says the Big Bang was the beginning of something smaller than the universe. IV - Student says the Big Bang is an event that happened to something smaller than the universe. V - Student says the Big Bang is the beginning of space. VI - Student says the Big Bang is the beginning of time. VII - Student talks about the creation/production of elements. VIII - Student says the Big Bang was an explosion. IX - Student says the Big Bang was not an explosion. X - Student says matter existed before the Big Bang. XI - Student says there was a dense piece of matter before the Big Bang. XII - Student talks about matter coming together before the Big Bang. XIII - Student says matter formed from energy. XIV - Student says the early universe was hot, dense, and/or small. XV - Student says the Big Bang was an event that happened in empty space. XVI - Student gives irrelevant information. XVII - Student gives some other reason not specified above. XVIII - Student says s/he has no idea. XIX - Answer field is blank or the student provided no reason.
2	Students in category 2 talk about the Big Bang using at least one of the following elements: • It's the beginning of the universe, space, and/or time. • It's when (some) of the elements were created. • It refers to how the universe was once hot, small, and dense. Students in this category either provide no further information or they use one of the following elements: • The Big Bang was an explosion. • Matter existed before the Big Bang. • The Big Bang was an event that happened in empty space. Additionally, any student whose ideas about the universe do not fit into categories 3 or 1 should be placed here.	
1	Students in category 1 talk about the Big Bang as the beginning of something smaller than the universe or and event that happened to something smaller than the universe.	
0	Students in category 0 write nothing (the answer field is blank), or they provide information that doesn't answer the question, or they say they have no idea and provide no further information.	

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response with contain every element in a given score category, you

must also list the reasoning elements the student uses.

Example responses and scores:

"The Big Bang Theory, in a nutshell, says that the universe started out as very hot + dense, and over time that energy was converted into matter as the distance between matter increased and as the universe expanded."

> **How this response is scored:** 3_XIV_XIII_II. This student gets and overall score of 3 because in her response she mentioned expansion (II), the formation of matter from energy (XIII), and how the universe was once hot and dense (XIV).

"Big Bang Theory (not the TV show) means that all the stuff, matter, energy, whatever that exists in the universe now was once scrunched up into a teeny tiny little thing which then exploded and expanded rapidly."

> **How this response is scored:** 2_X_XI_II_VIII. This student receives a 2 for this response because she said the Big Bang is the beginning of expansion (II), a dense piece of matter existed before the Big Bang (X and XI), and the Big Bang was an explosion (VIII).

"'the big bang theory' is when an asteroid that was headed toward earth struck the earth and every thing that was alive died --- then as time went on things started growing and living again."

> **How this response is scored:** 1_IV. This response earns an overall score of 1 because it talks about the Big Bang as an event that happened to something smaller than the universe (IV).

"Big Bang theory means an idea or a way of critical thinking between astronomers, and/or scientists to the way they see the earth and universe through telescope or even without one through basic observation."

> **How this response is scored:** 0_XVI. This response does not really answer the question and gives a lot of irrelevant information (XVI).

Item 3: Each dot in the picture on the left is a galaxy. The Milky Way Galaxy (the one we live in) is at the center of the picture. All of the galaxies inside the circle can be seen from Earth. Any galaxies that exist outside the circle are so far away that their light has not had time to reach Earth. Describe what inhabitants of Galaxy X probably see when they look in the direction of the arrow.

> **What this item measures:** This item looks at whether students realize there are galaxies beyond our observable universe and that each location should have its own observable universe that looks like any other observable universe.

Scoring rubric for item 3.

Score	Response Includes	Reasoning Elements
2	- inhabitants of Galaxy X see something similar to what we see or they see more stars and/or galaxies	I - Student says they'd see more galaxies. II - Student says they'd see more stars. III - Student says they'd see more planets. IV - Student says they'd see more "objects" (otherwise unspecified) V - Student says they'd see something similar to what we see. VI - Student says they'd see things we cannot see. VII - Student says they'd see nothing/blackness/empty space. VIII - Student says they'd see objects separated by greater distances than what we see. IX - Students say Galaxy X's observable universe extends into regions outside of our observable universe. X - Student gives irrelevant information. XI - Student gives some other reason not specified above. XII - Student has no idea. XIII - Answer field is blank.
1	- does not specify if inhabitants of Galaxy X see something similar to what we see or they see more stars and/or galaxies - says the universe outside our observable universe is not similar to our observable universe	
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	

Example responses and scores:

"Galaxy X probably sees many galaxies in that direction. They probably look just as surrounded as Earth."

> **How this response is scored:** 2_I_V. This student says Galaxy X would see more galaxies (I) and their overall view would probably be similar to our view (V).

"They probably see nothing since there is no light able to reach there."

> **How this response is scored:** 1_VII. This student gets a score of 1 because she says

inhabitants of Galaxy X would see nothing (VII).

Item 4: Circle the phrase that best completes the sentence. Surrounding the event called the BB was ____.

a) a region of space that includes nothing (empty space)
b) a region of space that includes particles and matter
c) neither a nor b

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

What this item measures: This item probes whether or not students think matter and space existed before the Big Bang.

Scoring rubric for item 4.

Score	Response Includes	Reasoning Elements
3	- correct answer (C) - says there no matter/only energy and no space before the Big Bang	I - Student says there was "nothing" (unspecified) before Big Bang. II - Student says there was "something" (unspecified) before the Big Bang. III - Student says there was no time before the Big Bang. IV - Student says there was time before the Big Bang. V - Student says there was no space before the Big Bang. VI - Student says there was space before the Big Bang. VII - Student says there was no matter before the Big Bang. VIII - Student says there was matter before the Big Bang. IX - Student says matter necessarily existed before the Big Bang. X - Students says matter before the Big Bang was all in one place/part of the Big Bang. XI - Student says space is never completely empty. XII - Student says there was only energy (no matter) at or before the Big Bang. XIII - Student says the Big Bang was the beginning of the universe/everything. XIV - Student says the Big Bang is an event that happened to something smaller than the universe. XV - Student says the Big Bang was an explosion. XVI - Student says we don't/can't know or describe what the universe was like this early. XVII - Student talks about what the universe was like after the Big Bang. XVIII - Student says this answer makes sense/the alternatives don't make sense. XIX - Student says this answer is better than the alternatives. XX - Student gives irrelevant information. XXI - Student gives some other reason not specified above. XXII - Student has no idea. XXIII - Answer field is blank.
2	- correct answer (C) - incorrect or incomplete reason	
1	- incorrect answer	
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	

Example responses and scores:

"C. Before the big Bang there was nothing. No particles, no matter annd no empty space."

> **How this response is scored:** 3_C_I_V_VII. This student picked the correct answer and defended her choice by saying there was nothing (I) before the Big Bang - including matter (VII) and empty space (V).

"C. there was nothing surrounding that matter"

> **How this response is scored:** 2_C_I_VIII. This student picked the correct answer (C), but does not get a 3 because he says there was pre-existing matter (VIII). He also says there was nothing before the Big Bang (I), but does not specify if there was empty space or not.

"A. There was no matter, nothing"

> **How this response is scored:** 1_A_I_VII. This student selects an incorrect answer (A) but says there was no matter (VII) and nothing (I) before the Big Bang.

Item 5: Independent of whether we know its true location, is there a center to the universe? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item examines whether or not students think the universe has a center and why.

Scoring rubric for item 5.

Score	Response Includes	Reasoning Elements
3	- rejects the idea of a center because the universe is infinite/has no edges - rejects the idea of a center because the universe looks the same from all directions	I - Student says the universe is infinite/has no edges. II - Student says the universe is the same everywhere. III - Student says the Sun is at the center. IV - Student says everything has a center/must have a center. V - Student says the center is where the Big Bang happened/where the universe began/where the universe is expanding from. VI - Student says things orbit the center of the universe. VII - Student says our ignorance prevents us from seeing the center. VIII - Student says technological limitations prevent us from seeing the center. IX - Student says there is no center or the center changes because of expansion/objects are in motion. X - Student says there is no center since the universe has no shape/an irregular shape/unknown shape. XI - Student gives irrelevant information. XII - Student gives some other reason not specified above. XIII - Student has no idea. XIV - Answer field is blank.
2	- rejects the idea of a center for a reason other than those listed in category 3	
1	- says universe has a center or unclear if idea of center rejected - says you can't see the center due to our ignorance/technological limitations - says the center changes with expansion	
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	

Example responses and scores:

"no I think the universe is this object with no center or edge it is constantly going"

> **How this response is scored:** 3_I. This student gets a 3 because she says the universe has no edges (I).

"no. Because it's a 4D universe it doesn't have a known shape but it does not have a center because we don't know where exactly the location of the universe is."

> **How this response is scored:** 2_X. This student denies a center, but defends his answer by saying the universe has an unknown shape (X).

"yes, because the expansion had to start from a center point"

> **How this response is scored:** 1_V. This student says there is a center and that's where expansion started from (V).

Item 6: If you could travel to any location in the universe, could you go to a place where there would be no galaxies in front of you? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes whether students think there is an edge to the universe (in the sense that there is an end to the distribution of galaxies).

Scoring rubric for item 6.

Score	Response Includes:
3	- no, there will always be galaxies around you
2	- yes, if you go to a black hole or some other unintended or not quite correct reason.
1	- yes, there are regions where distribution of galaxies peters out/where there are no galaxies
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.

NB: There are no reasoning elements listed since each score is associated with its own unique reasoning element.

Example responses and scores:
"No, there are galaxies everywhere."

> **How this response is scored:** This response falls into category 3 since the student says there will always be galaxies around you.

"Um...I suppose you could go to a black hole and stare into it. Black holes are really super dense and no light escapes them so you wouldn't be able to see galaxies."

> **How this response is scored:** This student gets a score of 2 because she talks about looking into a black hole.

"Yes that would be toward the outside of the universe."

> **How this response is scored:** This response earns a score of 1 because the student says there are regions with no galaxies.

Item 7: Which of the following statements (a - d) are true? Circle all that apply.
In general, the expansion of the universe causes the distances between ________.
a) planets in the solar system to increase.
b) stars in the galaxy to increase.
c) galaxies in the universe to increase.
Explain your reasoning for your choice(s).

> **What this item measures:** This item tests whether or not students realize that the expansion of the universe only affects distances between galaxies, not distances between stars within a galaxy or planets within a solar system.

Scoring rubric for item 7.	
Score	**Response Includes:**
2	- idea that expansion only affects the distances between galaxies.
1	- idea that expansion affects distances other than those on galactic scales or that we cannot tell.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.

NB: There are no reasoning elements listed since each score is associated with its own unique reasoning element.

Example responses and scores:

"C. Stars and planets move toward and away from each other in the universe."

> **How this response is scored:** This student gets an overall score of 2 because she says that expansion only affects the distances between galaxies.

"A, B. Planets + stars expand their orbits. I'm not sure if galaxies expand."

> **How this response is scored:** This student gets an overall score of 1 because she claims that the distances other than those between galaxies increase due to expansion.

Form C

Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize the temperature of the universe has cooled over time.

The scoring rubric for item 1 is on the next page.

Scoring rubric for item 1.

Score	Response Includes	Reasoning Elements
3	- claim that universe has cooled over time. - reason based on the expansion of the universe.	I - Student says the temperature increased. II - Student says the temperature decreased. III - Student says the temperature stayed the same.
2	- claim that universe has cooled over time. - no reasons given or incomplete or incorrect reasons.	IV - Student says the temperature changed, but doesn't specify if it went up or down. V - Student talks about expansion and/or how the density of the universe changed.
1	- any claim other than the universe has cooled over time.	VI - Student talks about the birth/formation of stars/a star.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	VII - Student talks about the birth/formation of planets/a planet. VIII - Student talks about the birth/formation of galaxies/a galaxy. IX - Student talks about the birth/formation of unspecified objects. X - Student talks about the death of stars. XI - Student talks about the death of planets. XII - Student talks about the death of galaxies. XIII - Student talks about the death of unspecified objects. XIV - Student talks about changes during the lives of stars/a star. XV - Student talks about changes during the lives of planets/a planet. XVI - Student talks about changes during the lives of galaxies/a galaxy. XVII - Student talks about changes during the lives of unspecified objects. XVIII - Student talks about how the universe is big. XIX - Student talks about competing effects canceling out. XX - Student gives irrelevant information. XXI - Student gives other reason not specified above. XXII - Student says s/he has no idea. XXIII - Response field left blank.

NB**:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"The temp. has gotten cooler. Again assuming no matter is created in the expansion process of the universe than while the universe has expanded it has become less dense and as you become less dense you cool off."

> **How this response is scored:** 3_II_V. This student gets a 3 because he says the temperature decreases (II) and he explains this in terms of the expansion of the universe (V).

"It could be getting colder as time goes on because as the universe expands there is more and more area for the sun and other giant stars to heat, and also the sun is continuously collapsing on itself (I think?) which might mean its getting smaller, just like other giant stars in the universe are which could also be making it cooler as time progresses."

> **How this response is scored:** 2_V_XVIII_XIV_II. This student correct says that the universe is cooling (II) and relates this to its expansion (V). However, she also relates the cooling to changes during the lives of stars (XIV) and mentions the big region (XVIII) stars have to heat. She thus receives an overall score of 2.

"It has to have changed b/c Galaxies are colliding, New stars are formed, and many stars die. The temperature can vary."

> **How this response is scored:** 1_IV_XVI_VI_X. This student says the temperature changed, but does not specify if he thinks it went up or down (IV) so he gets an overall score of 1. His answer also contains information on how galaxies change (XVI), the formation of stars (VI), and the death of stars.

Item 2: How does the total amount of matter in the universe right now compare to the total amount of matter in the universe at the very beginning of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize that the total amount of matter is not conserved in the universe and that the early universe was pure energy.

Scoring rubric for item 2.

Score	Response Includes	Reasoning Elements
3	- claim that there is more matter in the universe now. - reason based on the idea that the early universe was pure energy and/or contained no matter until the temperature dropped.	I - Student says there was no matter/only energy in the beginning. II - Student talks about the temperature cooling and/or that matter formed from energy. III - Student says there is more matter now because the universe is expanding. IV - Student says the amount of matter increases as objects form and/or evolve. V - Student says the amount of matter decreases as objects form and/or evolve. VI - Student says the amount of matter increases as objects interact and/or die. VII - Student says the amount of matter decreases as objects interact and/or die. VIII- Student says the amount of matter does not change. IX - Student gives irrelevant information. X - Student gives other reason not specified above. XI - Student says s/he has no idea. XII - Response field left blank or no reason given.
2	- claim that there is more matter in the universe now. - no reasons given or incomplete or incorrect reasons.	
1	- any claim other than there is more matter in the universe now.	
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	

***NB*: You may assign multiple reasoning elements to a single response.

Example responses and scores:

"There is tons more matter now. Then there was tons of energy, but now it has transformed into matter."

> **How this response is scored:** 3_I_II. This student says there used to be a lot of energy in the universe (I) but that some of it was subsequently transformed into matter (II), so she gets an overall score of 3.

"I think there is more because the universe is expanding and things are being created along w/expansion to 'fill it up'"

> **How this response is scored:** 2_III_IV. This student says the amount of matter in the universe has increased, but explains this in terms of the expansion of the universe (III) and new objects forming (IV). Therefore, she gets an overall score of 2.

"There is just as much matter now then then. It has just expanded."

> **How this response is scored:** 1_VIII. This student gets a 1 because he says the amount of matter hasn't changed (VIII).

Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize that the density of matter has decreased over time due to expansion..

Scoring rubric for Item 3.

Score	Response Includes	Reasoning Elements
3	- claim that the density has decreased over time - reason based on the universe becoming bigger over time	I - Student says the density increases. II - Student says the density decreases. III - Student says the density is constant.
2	- claim that the the density has decreased over time - no reasons given or incomplete or incorrect reasons	IV - Student says the density changes, but doesn't specify how. V - Students says the amount of matter in the universe has changed.
1	- any claim other than the density has decreased over time	VI - Student says the amount of matter in the universe has not changed.
0	- nothing (answer field left blank) - information that does not answer the question - no idea	VII - Student says the size of the universe has changed/stuff spread out. VIII - Student says the size of the universe has not changed/stuff hasn't spread out. IX - Student talks about objects forming over time. X - Student talks about objects evolving over time. XI - Student talks about objects dying/breaking up. XII - Student talks about the effects of gravity. XIII - Student talks about matter changing forms. XIV - Student gives some other reason not specified above. XV - Student gives irrelevant information. XVI - Student says s/he has no idea. XVII - Answer field is blank or no response given.

***NB*:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"It was more dense @ the beginning, and is now less dense b/c the universe has expanded."

> **How this response is scored:** 3_II_VII. This student gets an overall score of 3 because she correctly said that the density has dropped (II) due to the expansion of the universe (VII).

"I think it's got less dense I mean there was none, then the Big Bang and now it formed."

> **How this response is scored:** 2_II_V_IX. This student gets an overall score of II because, although he said the density decreased (II), he also talked about how the amount of matter has changed (V) and objects forming (IX).

"It has changed, I couldn't say if its gotten larger or smaller but since D=M/V, and the mass has not changed, but I'm pretty sure V has, the D would be different."

> **How this response is scored:** 1_IV_VI_VII. This student says the density changes, but can't say how (IV), so she gets an overall score of 1. She also talks about how the amount of matter has not changed (VI), how the size of the universe has changed (VII).

Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y?

a) less than 8 billion light-years apart b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether a student can relate distances and lookback times in an expanding universe.

Scoring rubric for item 4.

Score	Response Includes	Answer Choice	Reasoning Elements
3	- right answer (A) - correct reason	A - Student chooses answer A.	I - Student talks about the expanding universe/galaxies getting farther apart.
2	- right answer (A) - no reasons given or incomplete or incorrect reasons		II - Student talks about galaxies getting closer together.
		B - Student chooses answer B.	III - Student says light needs 8 billion years to travel 8 billion light-years.
1	- incorrect answer(s)		IV - Student says light-years are shorter than normal years.
0	- nothing (answer field left blank - information that does not answer the question - no idea	C - Student chooses answer C.	V - Student says light-years are longer than normal years.
		D - Student chooses answer D.	VI - Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.
		Z - Student makes no selection.	VII - Student says light may be slowed by something between the two galaxies.
			VIII - Student says we need more information on how galaxies moving apart.
			IX - Student says we need more information on how galaxies moving together.
			X- Student says 8 billion light-years is the minimum distance between Galaxies X and Y.
			XI - Student talks about time Galaxy X needs to see explosion.
			XII - Student says there's not enough information.
			XIII - Student gives irrelevant information.
			XIV - Student gives other reason not specified above.
			XV - Student says s/he has no idea.
			XVI - Response field left blank or no reason given.

Example responses and scores:

"A. The Galaxies are spreading apart so the light from the explosion had to catch up with the othe Galaxy which ws speeding away, so it took longer."

> **How this response is scored:** 3_A_I. The student chose the correct answer (A) and defended his answer by discussing the effects of expansion.

"A. it was in the galaxy so closer"

> **How this response is scored:** 2_A_XIV. This student selected the right answer (A) but gave a reason that does not match any of the common reasoning elements (XIV).

"B. The sight of the explosion travels @ the speed of light to Earth. 8-billion years @ the speed of light means 8-billion light years away."

> **How this response is scored:** 1_B_III. This student selected the wrong answer (B) and discussed how light would travel 8 billion light-years in 8 billion years.

Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode?

a) less than 8 billion light-years apart	b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart	d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether a student can relate distances and lookback times in an expanding universe.

Scoring rubric for item 5.

Score	Response Includes	Answer Choice	Reasoning Elements
3	- right answer (C) - correct reason	A - Student choose answer A.	I - Student talks about the expanding universe/galaxies getting farther apart.
2	- right answer (C) - no reasons given or incomplete or incorrect reasons		II - Student talks about galaxies getting closer together.
		B - Student chooses answer B.	III - Student says light needs 8 billion years to travel 8 billion light-years.
1	- incorrect answer(s)		IV - Student says light-years are shorter than normal years.
0	- nothing (answer field left blank - information that does not answer the question - no idea	C - Student chooses answer C.	V - Student says light-years are longer than normal years.
		D - Student chooses answer D.	VI - Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.
		Z - Student makes no selection.	VII - Student says light may be slowed by something between the two galaxies.
			VIII - Student says we need more information on how galaxies moving apart.
			IX - Student says we need more information on how galaxies moving together.
			X- Student says we need more information on the rate of expansion.
			XI - Student says answer depends on where, exactly, events happen in Galaxies X and Y.
			XII - Student says there's not enough information.
			XIII - Student gives irrelevant information.
			XIV - Student gives other reason not specified above.
			XV - Student says s/he has no idea.
			XVI - Response field left blank or no reason given.

Example responses and scores:

"C. The galaxies are drifting away and have been now for 8 B years."

> **How this response is scored:** 3_C_I. This student selected the correct answer (C) and explained his answer in terms of the expanding universe (I).

"C. The light you see is older. In the time light must travel, the light you are seeing might not be there

anymore"

> **How this response is scored:** 2_C_XIV. The student chose the right answer (C) but for a reason that does not match any of the listed reasoning elements (XIV).

"D. We do not know if the galaxies are moving toward or away from each other. The fact that the universe is expanding means it is more likely that the galaxies are farther apart from each other now than earlier. But there is no information pertaining to the direction of the galaxies in relation to each other."

> **How this response is scored:** 1_D_I_VIII_IX. This student talks about the universe expanding (I) but feels he needs more information on whether the galaxies are moving toward (IX) or away (VIII) from one another.

Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age?

a) less than 5 billion years old b) exactly 5 billion years old
c) more than 5 billion years old d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether a student can relate distances and lookback times in an expanding universe.

Scoring rubric for item 6.

Score	Response Includes	Answer Choice	Reasoning Elements
3	- right answer (B) - correct reason	A - Student choose answer A.	I - Student talks about the expanding universe/galaxies getting farther apart.
2	- right answer (B) - no reasons given or incomplete or incorrect reasons		II - Student talks about galaxies getting closer together.
		B - Student chooses answer B.	III - Student says light needs 8 billion years to travel 8 billion light-years.
1	- incorrect answer(s)		IV - Student says light-years are shorter than normal years.
0	- nothing (answer field left blank - information that does not answer the question - no idea	C - Student chooses answer C.	V - Student says light-years are longer than normal years.
		D - Student chooses answer D.	VI - Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.
		Z - Student makes no selection.	VII - Student says light may be slowed by something between the two galaxies.
			VIII - Student says we need more information on how galaxies moving apart.
			IX - Student says we need more information on how galaxies moving together.
			X- Student reasons that 13-8=5/light traveled for 8 billion years.
			XI - Student talks about what Galaxy Y is like after the explosion.
			XII - Student talks about the star exploding more than 8 billion years ago.
			XIII - Student talks about how it takes time to see events happen in the universe/lookback time.
			XIV - Student says expansion affects age.
			XV - Student says answer depends on where, exactly, events happen in Galaxies X and Y.
			XVI - Student says there's not enough information.
			XVII - Student gives irrelevant information.
			XVIII - Student gives other reason not specified above.
			XIX - Student says s/he has no idea.
			XX - Response field left blank or no reason given.

Example responses and scores:

"B. 13-8=5"

> **How this response is scored:** 3_B_X. This student used subtraction (X) to arrive at the correct answer (B).

"B. Different life forms could believe anything really. They could think its been 20 billion years. By guessing that it has been 5 billion years exactly would seem most logical."

> **How this response is scored:** 2_B_XVIII. This student chose the right answer (B) but for a reason that does not correspond to any of the above reasoning elements (XVIII).

"A. It's more than 8 billion L.Y. away so they are looking more than 8 billion years in the past."

> **How this response is scored:** 1_A_XII. This student said the explosion happened more than 8 billion years ago (XII) and thus chose the wrong answer (A).

Form D

Item 1: Which one graph best represents how objects actually move in our galaxy? ***Explain the reasoning behind your choice.***

> **What this item measures:** This item measures whether or not students can identify the correct rotation curve for the galaxy (Graph 2).

Scoring rubric for item 1.

Score	Response
2	- student selects graph 2
1	- student selects any graph other than graph 2
0	- nothing (answer field left blank) - information that does not answer the question - no idea

NB**:** Although the question asks for students to explain their reasoning, I did not find any patterns useful to highlight in the responses students gave. Most simply described the graph they chose. In the post-surveys, many mentioned "dark matter" but not at a level beyond declarative knowledge. This is probably okay, since subsequent items probe the students' understandings of rotation curves further.

Item 2: Based on your selection in Question 1, describe how the speeds at which stars A, B, and C orbit the galaxy compare to one another.

> **What this item measures:** This item measures whether students can use the rotation curve they selected in item 1 to correctly compare the speeds of the three stars.

Scoring rubric for item 2.

Score	Response
2	- student correctly relates the speeds of the three stars based on the graph s/he chose
1	- student incorrectly relates the speeds of the three stars based on the graph s/he chose
0	- nothing (answer field left blank) - information that does not answer the question - no idea

Even though item 3 is not "officially" scored, still record a score of 0, 1, or 2. A student earns a 2 if s/he ranks the forces based on the idea that the velocity of a star directly corresponds to the gravitational force it feels (e.g. a student who says all stars move with the same velocity should say that all stars feel the same force, or a student who says stars closer to the galaxy's center are faster should likewise assign larger forces to stars closer to the center). Otherwise, any other reason is scored a 1. If there is no answer or the student say s/he has no idea with no further comment, give her/him a 0.

Item 4: Based on your previous answers, how is matter distributed in our galaxy? Pick the best answer from the following choices (a-c).

a) Most of the matter in the galaxy is located in the center.
b) Most of the matter in the galaxy is located in the center and spiral arms.
c) Neither a nor b.

Explain your reasoning for your choice.

> **What this item measures:** This item measures whether students realize that matter is not concentrated in the center of the galaxy.

Scoring rubric for item 4.

Score	Response Includes	Reasoning Elements
3	- correct answer (C). - talks about dark matter or how there's more matter than we can see throughout the galaxy or how light doesn't accurately predict where matter is located.	I - Student talks about gravity. II - Student talks about unspecified forces or pulls. III - Student talks about spinning or revolving. IV - Student talks about visible/normal matter. V - Student talks about unseen/dark matter.
2	- correct answer (C) with incomplete or incorrect reasoning. - incorrect answer, but one that is consistent with the student's answers to previous questions - incorrect answer, but reason includes a discussion about dark matter.	VI - Student talks about the galaxy's central black hole. VII - Student talks about stellar evolution. VIII - Student talks about galactic evolution. IX - Student talks about the center/bulge. X - Student talks about the spiral arms/disk. XI - Student talks about the halo/what surrounds the galaxy.
1	- incorrect answer with incorrect reasons	XII - Student talks about an even or equal distribution of matter.
0	- nothing (answer field left blank) - information that does not answer the question - no idea	XIII - Student relates the location(s) of mass to observations or previous answers. XIV - Student talks about where the galaxy appears densest or brightest. XV - Student gives irrelevant information. XVI - Student gives other reason not specified above. XVII - Student says s/he has no idea. XVIII - Response field left blank.

Example responses and scores:

NB: All of these responses are taken from the same student.

Item 1: *"3. Objects closest to the center of the galaxy move fastest because they are receiving the most gravity and have to move fast enough so that they don't get pulled to the center"*

How this response is scored: The student chose the wrong rotation curve, so he receives a 1.

Item 2: *"Star A is moving the fastest, then B, then C. Again, the closer the object to the center where gravity is pulling, the objects must be fast to avoid being pulled in."*

How this response is scored: This answer is consistent with the graph the student selected in item 1, so he receives a 2 on this question.

Item 3: *"The net g force on A would be greatest because it is closest to the center. star B would be next. then Star C. The net gravitational force is dependent on its distance from the center"*

How this response is scored: This answer is consistent with the student's previous answers, so he receives a 2.

Item 4: *"The majority of matter is located at the center of the galaxy where the gravitational force is strongest due to the high mass of objects such as a black hole?"*

How this response is scored: 2_IX_I_VI. This answer is consistent with the student's previous answers, so he receives a 2. He talks about the center of the galaxy (IX), gravity (I), and a black hole (VI) so he does not earn a higher score.

Appendix G

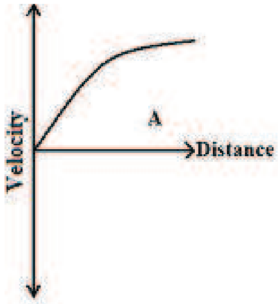
Fall 2010 Surveys

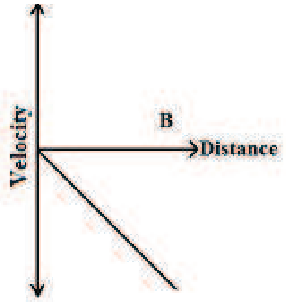
G.1 Form A

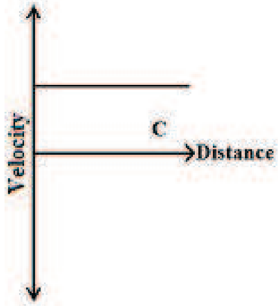
The fall 2010 version of Form A begins on the next page.

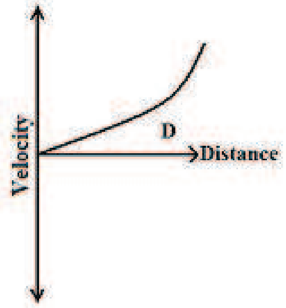
Name: ________________________________ **Conceptual Cosmology Survey - Form A**

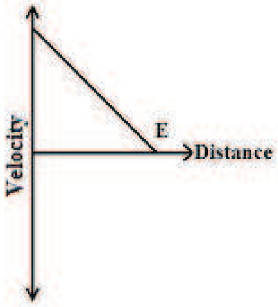
Below are eight graphs (A - H) showing how fast galaxies are moving (velocity) versus their distances away from us.

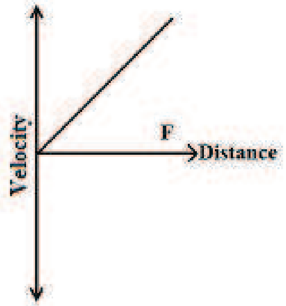


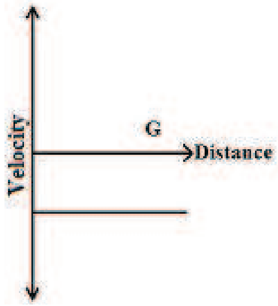


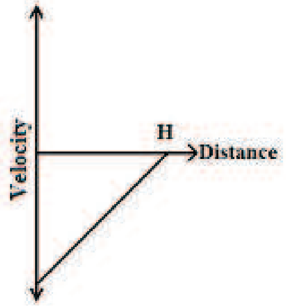












1) Which graph or graphs (A-H), if any, show a universe that is expanding at a constant rate? Explain your reasoning for your selection(s). If your answer is "none," explain why.

2) Which graph or graphs (A-H), if any, show a universe that is contracting at a constant rate? Explain your reasoning for your selection(s). If your answer is "none," explain why.

3) Which graph or graphs (A-H), if any, show a universe that is expanding at a faster and faster rate over time? Explain your reasoning for your selection(s). If your answer is "none," explain why.

4) Which graph or graphs (A-H), if any, show a universe that is expanding at a slower and slower rate over time? Explain your reasoning for your selection(s). If your answer is "none," explain why.

Figure 1, at right, is a possible graph showing how fast galaxies move away from us in the expanding universe.

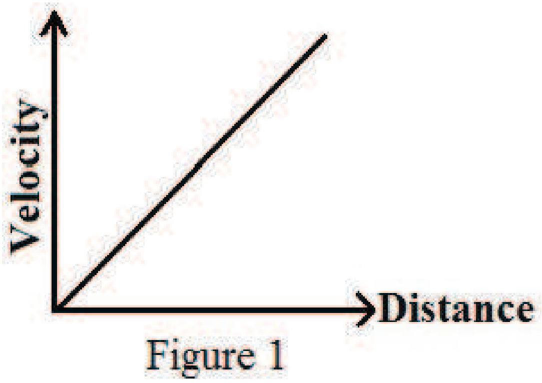
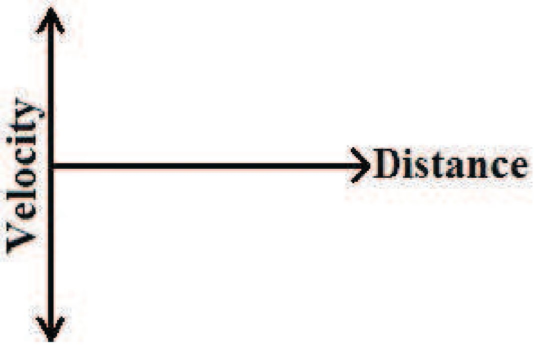


Figure 1

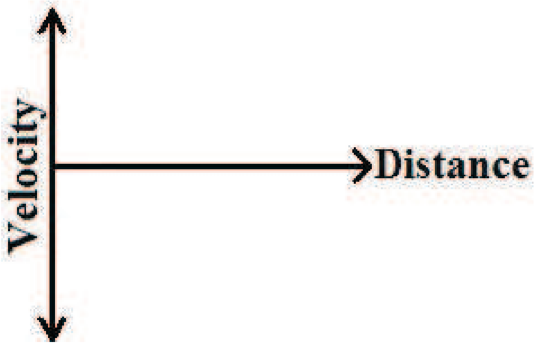
5) Use the blank graph provided below to draw what you think Figure 1 would look like if the universe had been expanding twice as fast. Explain the reasoning behind the graph you drew.

If you don't have enough information to do this, explain what else you need to know.



6) Use the blank graph provided below to draw what you think Figure 1 would look like for our universe if it took much longer to reach its current size. Explain the reasoning behind the graph you drew.

If you don't have enough information to do this, explain what else you need to know.



7) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

8) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

G.2 Form B

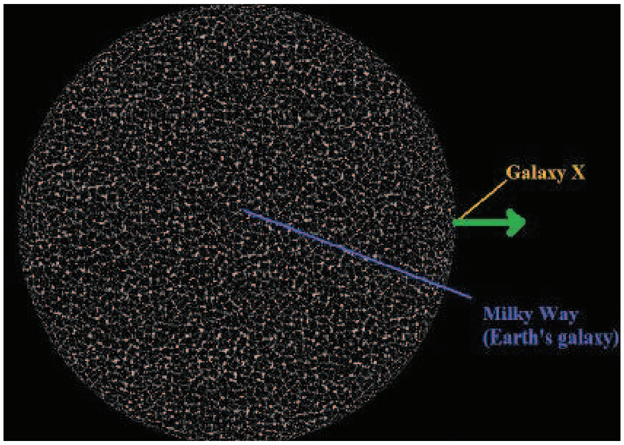
The fall 2010 version of Form B begins on the next page.

Name: ______________________________________ **Conceptual Cosmology Survey - Form B**

1) Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." Provide a drawing if possible to help illustrate your thinking.

2) Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory." Provide a drawing if possible to help illustrate your thinking.

3) Each white dot in the picture on the right is a galaxy. The Milky Way Galaxy (the one we live in) is at the center of the picture. All of the galaxies inside the circle can be seen from Earth. Any galaxies that exist outside the circle are so far away that their light has not had time to reach Earth. Describe what inhabitants of Galaxy X probably see when they look in the direction of the arrow.



4) Circle the sentence that best describes the universe at the time of the Big Bang:

a) In the beginning, there was space in the universe surrounding the location of the Big Bang but this space was empty of all matter.

b) In the beginning, there was space in the universe surrounding the location of the Big Bang and matter already existed in this space.

c) I think of the Big Bang differently than a or b.

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

5) Independent of whether we know its true location, is there a center to the universe? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

6) If you could travel to any location in the universe, could you go to a place where there are no galaxies in front of you (a.k.a. empty space)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

7) Which of the following statements (a - d) are true? Circle all that apply.
In general, the expansion of the universe causes ________.
a) the distances between <u>planets in the solar system</u> to increase.
b) the distances between <u>stars in the galaxy</u> to increase.
c) the distances between <u>galaxies in the universe</u> to increase.
d) None of the above.
Explain your reasoning for your choice(s).

8) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

9) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

G.3 Form C

The fall 2010 version of Form C begins on the next page.

Name: ___________________________________ **Conceptual Cosmology Survey - Form C**

1) Over time, would you say the <u>temperature</u> of the universe has increased, decreased, or stayed the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

2) How does the <u>total amount of matter</u> (not energy) in the universe *right now* compare to the total amount of matter at the *very beginning* of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

3) Over time, would you say the overall <u>density of matter</u> in the universe has increased, decreased, or stayed the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

<u>Questions 4-6 refer to this situation:</u> Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took exactly 8 billion years to reach Galaxy X.

4) How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y?
a) less than 8 billion light-years apart b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

5) How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode?
a) less than 8 billion light-years apart b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

6) The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age?

a) less than 5 billion years old b) exactly 5 billion years old
c) more than 5 billion years old d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

7) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

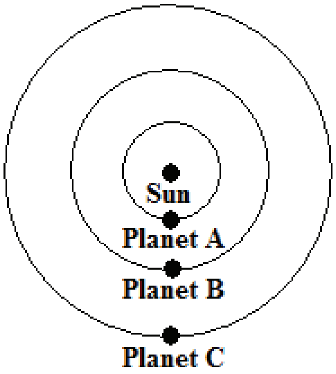
8) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

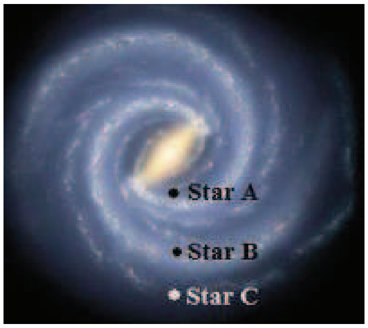
G.4 Form D

The fall 2010 version of Form D begins on the next page.

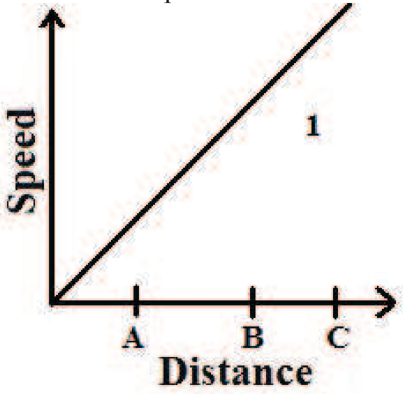
Name: ______________________________ **Conceptual Cosmology Survey - Form D**

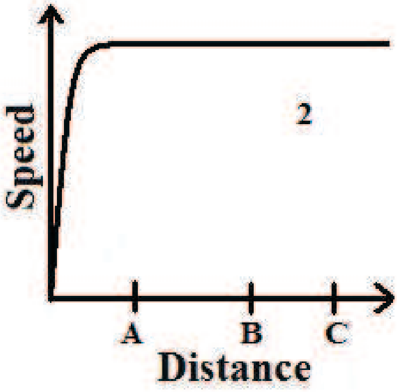
On the left is a picture of a solar system with three planets (A - C). On the right is a picture of a spiral galaxy which shows the locations of three stars (A - C).

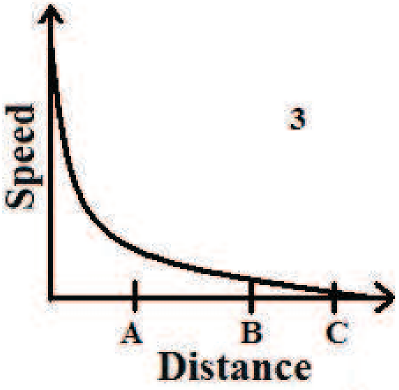


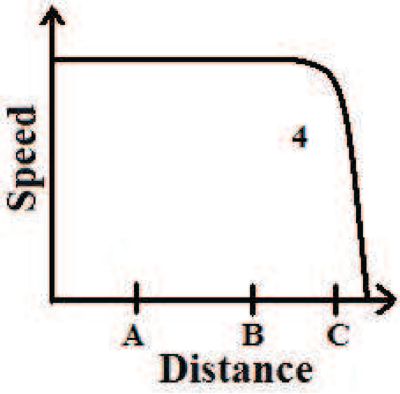


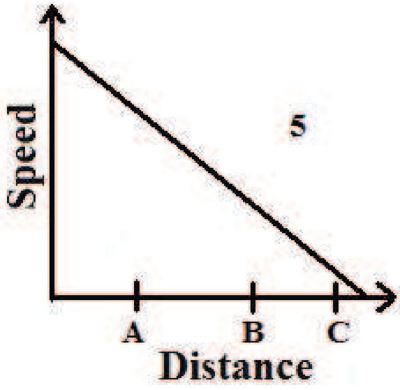
Astronomers make plots called rotation curves when they observe how fast planets orbit the Sun versus their distance from the Sun. They also plot rotation curves when they observe how fast stars orbit the center of a galaxy versus their distance from the center of the galaxy. Below are some possible rotation curves (1-6).

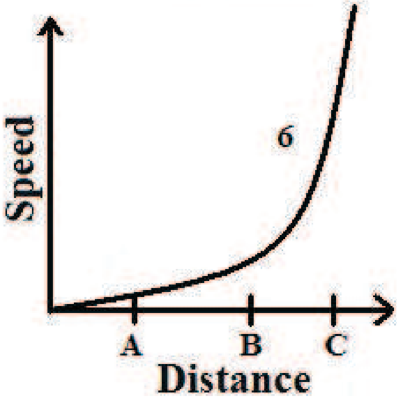












1) Which graph (1-6) best represents how **<u>planets</u>** orbit the Sun?

2) Which graph (1-6) best represents how **<u>stars</u>** orbit the center of the galaxy?

3) Rank the speeds at which **Planets** A, B, and C orbit the Sun:
Ranking Order: Fastest speed 1________ 2________ 3________ Slowest
Or, all the planets orbit at approximately the same speed. _____ (indicate with a check mark)
Explain your reason for ranking this way:

4) Rank the speeds at which **Stars** A, B, and C orbit the galaxy.
Ranking Order: Fastest speed 1________ 2________ 3________ Slowest
Or, all the stars orbit at approximately the same speed. _____ (indicate with a check mark)
Explain your reason for ranking this way:

5) Based on your previous answers, how is matter distributed in **solar systems**? Pick the best answer from the following choices (a-c).
a) Most of the matter in the solar system is located in the Sun.
b) Most of the matter in the solar system is evenly distributed throughout the Sun and planets.
c) My thinking is different than a and b.
Explain your reasoning.

6) Based on your previous answers, how is matter distributed in **spiral galaxies**? Pick the best answer from the following choices (a-c).
a) Most of the matter in the galaxy is located in the center.
b) Most of the matter in the galaxy is located in the spiral arms.
c) My thinking is different than a and b.
Explain your reasoning.

7) Based on your answers for Questions 1-7, do **stars** orbiting the center of a galaxy act like **planets** orbiting the Sun? If yes, explain why. If no, explain why not.

8) Did you take a class in **high school** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no
9) ***Not including this class***, did you take a class in **college** that covered the expansion of the universe? *(Check either "yes" or "no".)* ____ yes ____ no

Appendix H

Fall 2010 Lecture-Tutorials

H.1 Hubble's Law

The fall 2010 version of the "Hubble's Law" lecture-tutorial begins on the next page.

Hubble's Law

Part I: Expansion, Distance, and Velocity

Consider the small section of the universe containing four galaxies (A-D), shown in Figure 1 below. The distances between each galaxy are also shown.

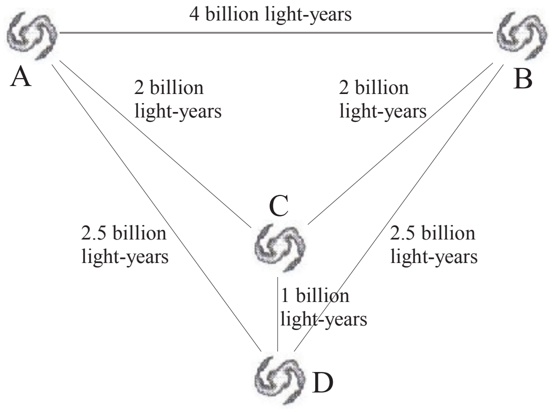


Figure 1

1) Imagine that this section of the universe doubles in size over time due to the expansion of the universe. Draw what the above section of the universe would look like after it doubles in size. Be sure to identify the new distances between the galaxies.

Hubble's Law

2) Which of the galaxies (B-D) increased its distance from Galaxy A by the greatest number of light-years during this time? Explain your reasoning.

3) Two students are discussing their answers to Question 2:

Student 1: *All of the distances doubled, so all of the distances increased by the same amount. There is no one galaxy whose distance from Galaxy A increased the most.*

Student 2: *You're right that all the distances double in size, but I don't agree that they all increase by the same number of light-years. Since Galaxy B was the farthest away from Galaxy A initially, its distance will increase by the greatest number of light-years when this section of the universe doubles in size.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

4) Describe the relationship between a galaxy's distance from Galaxy A and the speed at which that galaxy appears to be moving away from Galaxy A.

5) Is the relationship you described in Question 4 unique to Galaxy A, or would you observe the same relationship (between distance and speed) if you lived in one of the other galaxies? Explain your reasoning.

Hubble's Law

Part II: Understanding Hubble's Law and Hubble Plots

The relationship you described in Questions 4 and 5 is called *Hubble's law*. We can depict Hubble's law with the graph shown at right. This graph plots the speed at which a galaxy appears to move away from us versus its distance from us. This type of graph is called a *Hubble plot*. Each dot on the plot represents a different galaxy.

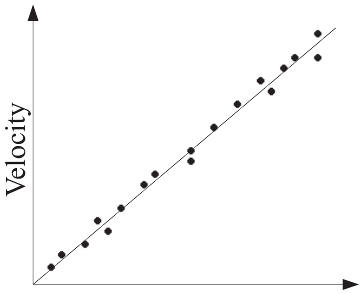


Figure 2

6) Explain how the Hubble plot shown in Figure 2 above is consistent with the relationship you described in Question 4.

7) Imagine the Hubble plot shown in Figure 2 represents a universe that doubles in size over a certain amount of time. Which of the Hubble plots shown in Figures 3 and 4 below might represent a universe that triples in size over the same amount of time? Explain your reasoning.

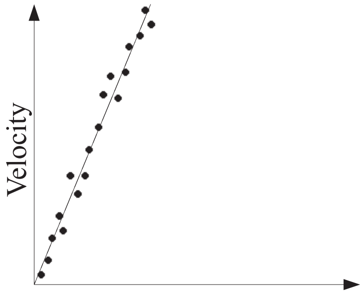


Figure 3

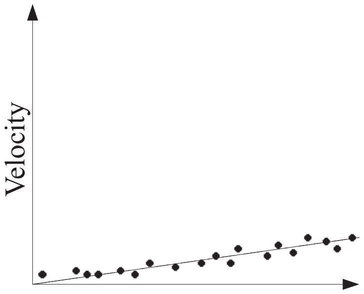


Figure 4

The *expansion rate* of the universe determines how fast it is growing in size. For example, a universe that is tripling in size has a faster *expansion rate* than a universe that is doubling in size over the same amount of time. In a Hubble Plot, the *expansion rate* is indicated by the slope of the graph. A steep slope indicates a fast expansion rate, while a flat slope indicates a slow expansion rate.

8) Would you say the expansion rate for the universe represented in Figure 2 is constant, increasing, or decreasing with time?

Hubble's Law

9) Rank (from fastest to slowest) the expansion rates of the three different universes represented in Figures 2, 3, and 4. Explain your reasoning.

10) If the expansion rate of our universe had been faster, would the universe have reached its current size earlier in its history or later? Explain your reasoning.

11) Complete the sentence below using the words provided in parentheses ().

For two universes that are the same size, the universe with the faster expansion rate must be __________(younger/older) than the universe with the slower expansion rate.

12) If the above Hubble plots (Figures 2-4) represent three universes that are the same size, which Hubble plot belongs to the youngest universe? Explain your reasoning.

We can imagine many different Hubble plots, which may or may not represent how galaxies move as a result of expansion.

13) On the blank graph in Figure 5 below, draw a Hubble plot for which the expansion rate is zero.

14) On the blank graph in Figure 6 below, draw a Hubble plot for which the expansion rate is increasing throughout the lifetime of the universe.

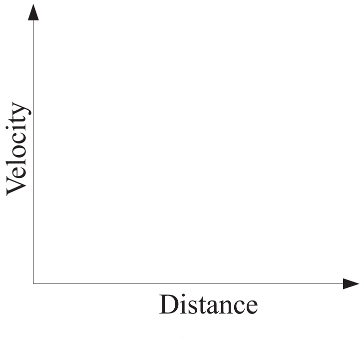


Figure 5

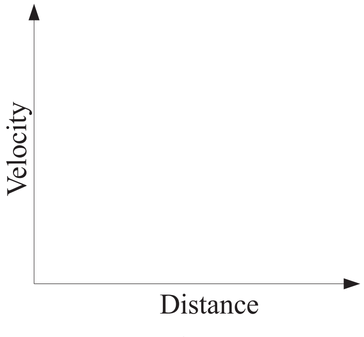


Figure 6

Hubble's Law

Part III: Our Universe

Recent observations indicate the Hubble plot for our universe actually looks more like the plot in Figure 7.

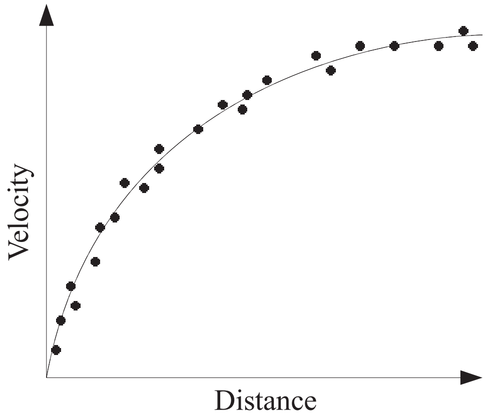


Figure 7

15) On Figure 7, draw a circle around the galaxies from which we receive information closest to our present time.

16) On Figure 7, draw a square around the galaxies from which we receive information furthest from our present time.

17) On Figure 7, write the letter A by the galaxies that are moving away from us with the fastest velocities.

18) On Figure 7, write the letter B by the galaxies that are moving away from us with the slowest velocities.

19) On Figure 7, write the letter S where the graph has the steepest slope.

20) On Figure 7, write the letter F where the graph has the flattest slope.

21) On Figure 7, write the Greek letter α by the portion of the graph that corresponds with the fastest expansion rate.

22) On Figure 7, write the Greek letter β by the portion of the graph that corresponds with the slowest expansion rate.

Hubble's Law

23) Based on the Hubble plot shown in Figure 7, would you say that the expansion rate of the universe is constant or changing with time? Explain your reasoning.

24) Based on the Hubble plot in Figure 7, is the expansion rate represented by the motion of galaxies far away from us faster than, slower than, or the same as the expansion rate represented by the motions of nearby galaxies? Explain your reasoning.

25) Based on the Hubble plot in Figure 7, is the expansion rate of the universe increasing or decreasing as time goes on? Explain your reasoning.

26) Consider the following debate between two students regarding their answer to the previous question:

Student 1: *The expansion rate of our universe must be slowing down as time goes on. If you look at the Hubble plot, you can see that the graph gets flatter. That means the farther away you look, the slower the expansion rate is. The rate at which the most distant galaxies are moving away from us has started to slow down and eventually the expansion rate of nearby galaxies will also slow down.*

Student 2: *I think you are reading the graph wrong. The slope of the graph tells you how fast the expansion rate of the universe is, not how fast a galaxy is moving. The farther we look into space, the further we are looking back in time. Since the slope of the Hubble plot is flatter in the past and steeper now, that means the expansion rate has sped up over time.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

27) Based upon your previous answers, is the graph you drew in Question 14 correct or does it need to be redrawn? Explain your reasoning.

Part IV: The Future of Our Universe

Far away galaxies appear dimmer in our sky than nearby galaxies. This is because we receive less light from far away galaxies than we do from nearby galaxies. So the farther away a galaxy is from us, the harder it is to see.

28) Complete the sentence below using the words provided in parentheses ().

Our observable universe is getting __________ *(darker/brighter) over time.*

Explain your reasoning.

29) Has the rate at which the universe is getting darker or brighter been the same for the entire age of the universe or is it getting darker at a faster (or slower) rate now as compared to the past? Explain your reasoning.

30) Is the above answer true for all locations in the universe? Explain your reasoning.

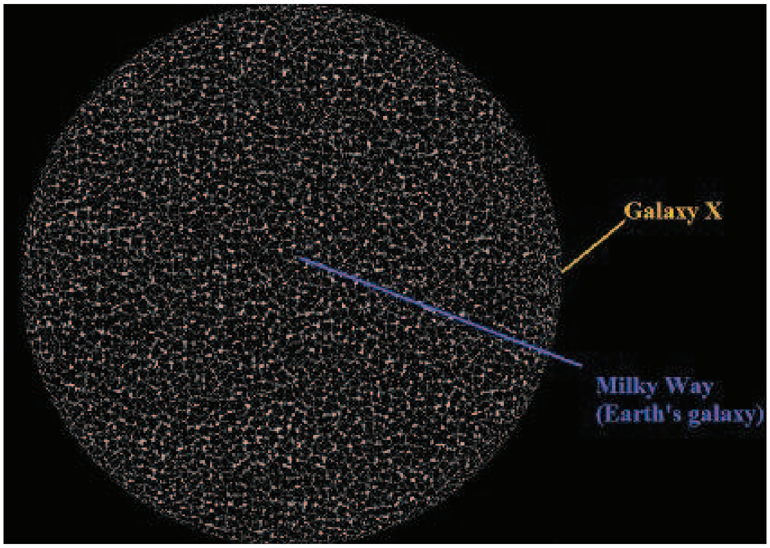
H.2 Making Sense of the Universe and Expansion

The fall 2010 version of the “Making Sense of the Universe and Expansion” lecture-tutorial begins on the next page.

Making Sense of the Universe and Expansion

Part I: The Observable Universe

Each dot in the picture below represents a galaxy. The Milky Way galaxy is represented by the dot at the center of the picture. All of the galaxies inside the circle can be seen from Earth. The circumference of this circle defines what is called our *observable universe*. Any galaxy that exists outside the circle is so far away that its light has not had time to reach Earth and is therefore not part of our observable universe.



1) Do you think the galaxies we can see from Earth are the only galaxies in the *entire universe*? Explain your reasoning.

2) Draw a circle around Galaxy X that represents its *observable universe*.

3) Is the *observable universe* that you drew for Galaxy X different in size than the *observable universe* for the Earth? Explain your reasoning.

Making Sense of the Universe and Expansion

4) Two students are talking about the *observable universe* for Galaxy X:

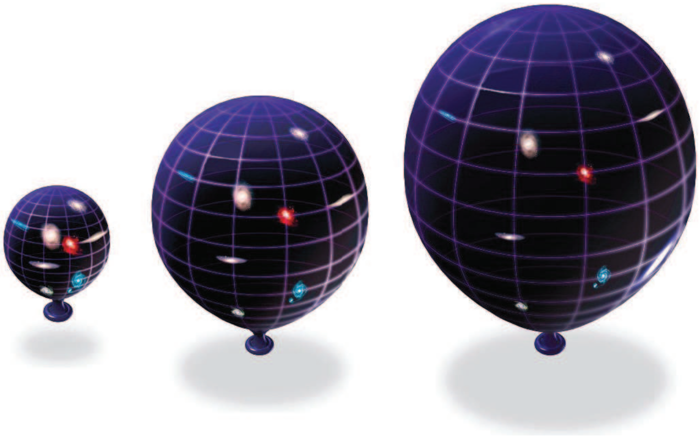
Student 1: *People living in Galaxy X have a strange view of the universe. When they look in one direction they see a bunch of galaxies, but when they look in the other direction all they see is empty space. Galaxy X must be at the edge of the universe since there's nothing but black, empty space beyond it. We're lucky we live at the center since we can see galaxies all the way out to the edge of the universe, no matter where we look.*

Student 2: *I think you're wrong. People living in Galaxy X would probably see a bunch of galaxies in every direction they look, but they can see some galaxies that we can't, just like we can see galaxies they can't. The observable universe for any galaxy should look similar to ours. I don't think we are at the center of the universe and I don't think Galaxy X is at the edge either.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

Part II: An Analogy for Expansion

One way to try to understand and envision the expansion of the Universe is by creating analogies that model the different aspects of our real expanding Universe. One way to model the expanding Universe is to use a "balloon" analogy. In this analogy, the space and time of the universe are modeled by the "surface" or "skin" of an expanding balloon. The *entire universe* only exists on the surface of the balloon. Light can only travel on the surface of the balloon.



Making Sense of the Universe and Expansion

5) Do objects, light, or events in the *entire universe* also exist inside or outside of the balloon's surface in this analogy?

6) If you were to travel from galaxy to galaxy along the surface of the balloon, would you ever encounter an edge?

7) If you were to travel over the entire surface of the balloon universe, would you ever find a location that you would consider to be the center of the *entire universe*?

8) Consider the following debate between two students about their answers to the previous questions:

Student 1: *Someone living on the surface of this balloon universe will definitely encounter an edge and a center. All they have to do is look from their location across the inside of the balloon to a location on the other side. The center of the inside of the balloon is the center of the universe, and the far side would be the edge of what they could see. So there's definitely a center and an edge to the universe in the balloon analogy.*

Student 2: *I think you misunderstand the analogy. The surface of the balloon is supposed to be the entire universe. The inside of the balloon isn't part of the universe and doesn't actually exist. You can't look through the inside of the balloon to the other side so there is no center in the middle or edge on the other side. In this analogy, people living in the balloon universe would never encounter a center or an edge.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

9) The balloons on page 2 represent the universe at different times during its history. Draw an arrow underneath the balloons that points from the earliest depiction of the universe to the latest.

Making Sense of the Universe and Expansion

10) Imagine you lived in a galaxy on the surface of the balloon. As the balloon expands, would all the other galaxies appear to move toward you or away from you?

11) Would your answers to the previous question be the same regardless of the galaxy in which you live, or would it change depending on the galaxy you inhabit?

12) In this analogy, do galaxies move relative to one another because they are traveling across the surface of the balloon, or do they move relative to one another because the balloon is expanding?

13) The balloon analogy is a helpful way to think about expansion, but no analogy is perfect. Some aspects of the real universe are captured by this analogy while others are not. The evidence we now have about the real universe implies the following statements (a-f) are all true. For each of these statements, state whether it is accurately captured by the balloon analogy or not, and explain your reasoning.

a) The real universe has no center.

b) The real universe has no edge.

c) The real universe is expanding.

d) The real universe is not round.

e) The real universe's expansion does not cause galaxies to change size.

f) The real universe is 4-dimensional (3 dimensions of space and 1 of time).

H.3 Expansion, Lookback Times, and Distances

The fall 2010 version of the “Expansion, Lookback Times, and Distances” lecture-tutorial begins on the next page.

Expansion, Lookback Times, and Distances

When the universe was 4 billion years old, Galaxy A was 3 billion light-years away from Galaxy B, as shown below. Imagine that the universe was not expanding, so the distance between Galaxy A and Galaxy B would not change over time.

1) A star explodes in Galaxy B producing a large amount of light. How long will the light from this explosion take to reach Galaxy A?

2) How far did the light travel on its journey to Galaxy A?

3) How old will the universe be by the time the light from the explosion reaches Galaxy A?

Because light takes time to travel from place to place in the universe, when we look at the night sky we are seeing stars and galaxies as they appeared in the past. For example, if we see a galaxy 1 million light-years away, we are seeing what the galaxy looked like 1 million years ago. We would say this galaxy has a lookback time of 1 million years. Lookback time is the amount of time light takes to travel to us from a distant object.

4) What is the lookback time inhabitants of Galaxy A associate with Galaxy B when they see the light from the explosion?

The real universe is expanding. This means the distance between galaxies is constantly increasing. Imagine that Galaxy A and Galaxy B are in an expanding universe.

5) While the light from the explosion is traveling from Galaxy B to Galaxy A, does the distance between the two galaxies stay the same, become larger, or become smaller?

Expansion, Lookback Times, and Distances

6) By the time the light from the explosion in Galaxy B reaches Galaxy A, is the distance between the galaxies more than, less than, or exactly 3 billion light-years?

7) By the time the light from the explosion in Galaxy B reaches Galaxy A, has more than, less than, or exactly 3 billion years elapsed since the star exploded?

8) By the time the light from the explosion in Galaxy B reaches Galaxy A, will the total distance traveled by the light be more than, less than, or exactly 3 billion light-years?

9) When the inhabitants of Galaxy A see the light from the explosion in Galaxy B, are they looking at an event with a lookback time of more than, less than, or exactly 3 billion years?

10) In the space below provide a sketch that explains the reasoning behind your answers to questions (6-9).

Expansion, Lookback Times, and Distances

11) Consider the discussion between two students regarding their ideas about two distant galaxies in an expanding universe.

Student 1: *Let's say light takes 5 billion years to travel from one galaxy to another. This means the two galaxies were separated by 5 billion light-years when the light began its journey.*

Student 2: *If the light traveled for 5 billion years, then the distance between the two galaxies must have been less than 5 billion light-years when the light began its journey because the distances between galaxies are always increasing in the expanding universe.*

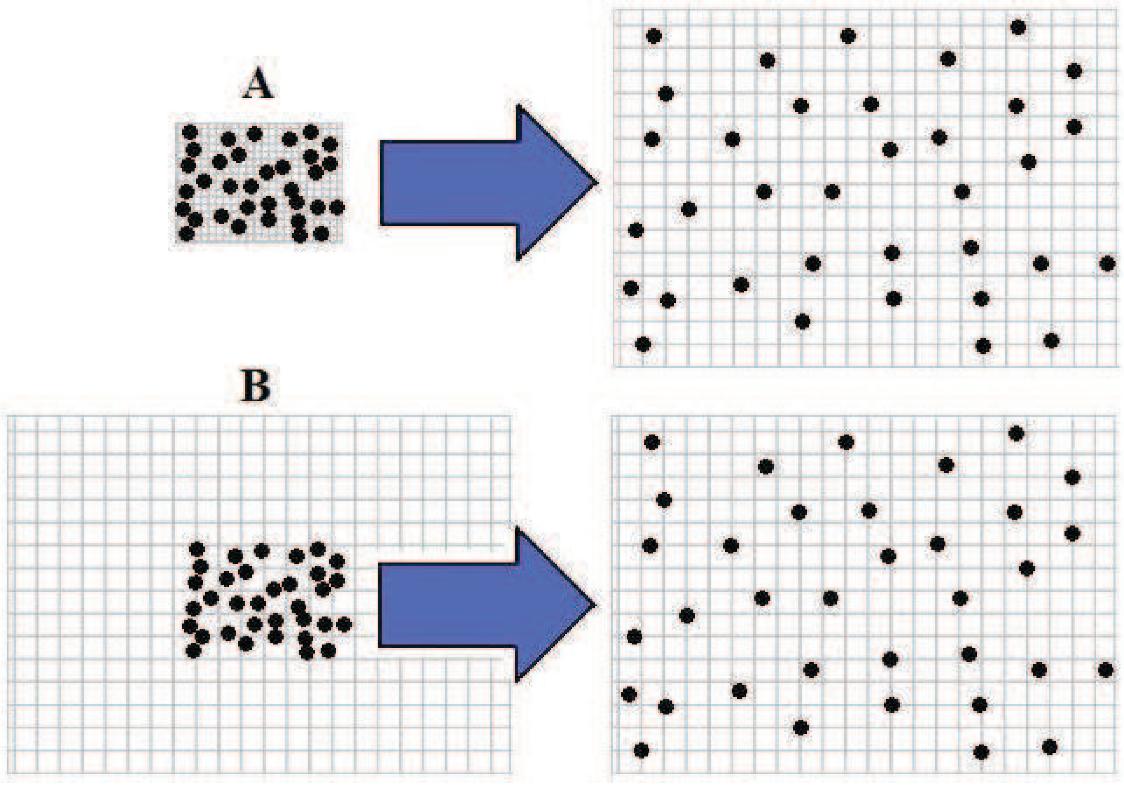
Do you agree with either or both of the students? Explain your reasoning.

H.4 The Big Bang

The fall 2010 version of “The Big Bang” lecture-tutorial begins on the next page.

The Big Bang

Consider the drawings provided below. Drawings A and B each represent a different way of thinking about how regions of the universe change over time. The dots in each drawing represent pieces of matter.



1) Which drawing, A or B is a better representation of the universe we observe? Explain your reasoning.

2) In Diagram A, is the universe becoming bigger, smaller, or staying the same size over time?

3) In Diagram B, is the universe becoming bigger, smaller, or staying the same size over time?

The Big Bang

4) Two students are debating their answers to Questions 2 and 3:

Student 1: *Both diagrams show the universe becoming bigger. In Diagram A, the grid has expanded and become larger. In Diagram B, the pieces of matter have spread out and take up a greater amount of space.*

Student 2: *I disagree. Only Diagram A shows the universe becoming bigger. In Diagram B the size of the grid doesn't change. The pieces of matter are just moving into an already existing empty space in a universe whose size doesn't change.*

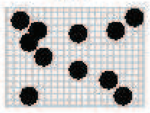
Do you agree or disagree with either or both of the students? Explain your reasoning.

5) Both drawings show the distance between matter increasing over time.
a) Which of the drawings shows this happening as the result of space expanding and which is a result of an outward explosion?

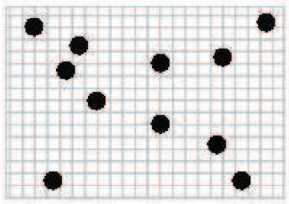
b) Which of the drawings is a more correct representation of our universe? Is your answer to this question consistent to your answer to Question 1? Explain your reasoning.

Consider the three drawings (C, D and E) shown below. These diagrams each represent a single region of the universe, but at different times during the history of the universe.

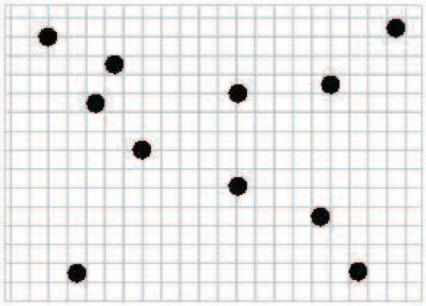
C



D



E



The Big Bang

6) Draw an arrow below drawings C, D, and E. The arrow should point from the drawing that represents the earliest time in the universe's history to the drawing that represents the latest time in the universe's history.

7) In which drawing does the region of space have:
a) the highest density?

b) the greatest concentration of energy?

c) the highest temperature?

Explain your reasoning.

8) Imagine you could watch the history of the universe like a movie playing backward. The movie starts today and ends at the beginning of the universe. Describe what you would see for every region of the universe as the movie played and you looked further back in time. Your answer should discuss how regions of the universe change in terms of temperature, and density, and size.

Your answers to the previous questions are all part of the *Big Bang Theory*. The Big Bang Theory does not say what the universe was like at the very first moment of time, which was about 13.7 billion years ago. It does, however, tell us how the universe changed after its first moment of existence.

The Big Bang

9) Three students are discussing their understandings of the Big Bang Theory:

Student 1: *I think I understand the Big Bang now. At the beginning, all the matter in the universe was compacted into a small, hot, dense ball. This ball of matter then exploded into empty space. When we look at the universe, we see galaxies moving away from us. The Big Bang model explains this, since all matter should be flying away from the center point of the explosion.*

Student 2: *I disagree. I think what the Big Bang Theory is saying is that all the matter in the universe was once compacted into a really dense and hot object that expanded over time. But there wasn't an explosion of matter into empty space. Instead, the universe carried pieces of matter away from each other as it expanded in size.*

Student 3: *You're both wrong. I agree that the universe was once smaller in size and that pieces of matter have been carried away from each other by the expansion of the universe. But remember how we learned from Einstein's equation* $E = mc^2$ *that matter can be converted into energy and energy can be converted into matter? I think this means that if we go back to the beginning of the universe, it would be so incredibly dense and hot that matter itself couldn't exist. I bet at the very beginning, the universe would have been composed of pure energy with no matter there at all.*

Which students do you agree or disagree with? Explain your reasoning.

10) Based on your previous answers, complete the following sentences:

The Big Bang Theory says that the universe started out with a/an __________ temperature and a/an ___________ density. Originally, there was no ___________, only pure ___________. From this initial state, each region of the universe ____________ in size. This caused its temperature and density to ___________. When the temperature was cool enough, energy could transform into ___________.

11) Look at drawing A again. Next to drawing A, make a drawing of what you think that region of the universe would have looked like at the very first instant it existed.

H.5 Dark Matter

The fall 2010 version of the “Dark Matter” lecture-tutorial begins on the next page.

Dark Matter

Part I: Motions of Planets

An object's orbit depends on the "mass inside" its orbit (also known as the *interior mass*). For a planet in our Solar System, you can find the interior mass by adding the Sun's mass to the mass of each object between the Sun and the planet's orbit. For example, the interior mass to Earth's orbit would be the Sun's mass plus the mass of Mercury plus the mass of Venus.

Here is a table that lists each planet, the mass inside each planet's orbit, and the speed at which the planets orbit the Sun.

Planet	Interior Mass (solar masses)	Orbital Speed (km/s)
Mercury	1.00	47.9
Venus	1.0000027	35.0
Earth	1.0000057	29.8
Mars	1.0000060	24.1
Jupiter	1.00096	13.1
Saturn	1.0012	9.66
Uranus	1.0013	6.81
Neptune	1.0013	5.43

1) Where is the vast majority of mass in the solar system located? What object or objects account for most of this mass?

2) Two students are discussing their answers to Question 1:

Student 1: *I think the majority of the mass in the solar system must include both the Sun and the planets. As you get farther away from the Sun, the interior mass gets bigger and bigger because you include more planets.*

Student 2: *I disagree. The majority of the mass in the solar system is from just the Sun by itself. Sure the mass gets a little bigger as you include more planets, but the extra mass is really small. Compared to the Sun, the planets hardly contribute any mass to the overall mass of the solar system.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

Dark Matter

3) How does the orbital speed of planets farther from the Sun compare to the orbital speed of planets closer to the Sun?

A planet's orbital speed depends on the gravitational force it feels. The strength of the gravitational force depends on the amount of mass that is inside of the planet's orbit as well as how far away the planet is from this interior mass.

4) How does the gravitational force on planets farther from the Sun compare to the gravitational force on planets closer to the Sun? Explain your reasoning.

5) Complete the blanks in the sentences of the following paragraph by either writing in the necessary information or circling the correct response. It may be helpful to base your responses on the information provided in the table above and your answers to the previous questions.

   There are ___ *planets inside Neptune's orbit and* ___ *planets inside Mercury's orbit. However, the interior mass for Neptune is* ___ (much greater than/approximately the same as/much less than) *the interior mass of Mercury. Neptune is* ___ (much closer to/much farther from/about the same distance from) *the Sun as/than Mercury. Therefore the gravitational force exerted on Neptune is* ___ (stronger/weaker/about the same strength) *as/than the force exerted on Mercury. As a result, Neptune has an orbital speed that is* ___ (much slower, much faster, about the same speed) *as/than the orbital speed of Mercury.*

If you could increase the amount of mass inside a planet's orbit, you would increase the gravitational force it would feel and thus increase the orbital speed of the planet. Imagine you were able to add a very, very large amount of mass distributed evenly *between* the orbits of Jupiter and Saturn.

6) Which planet(s) will experience an increase in gravitational force and an increase in orbital speed from this added mass? Explain your reasoning.

Dark Matter

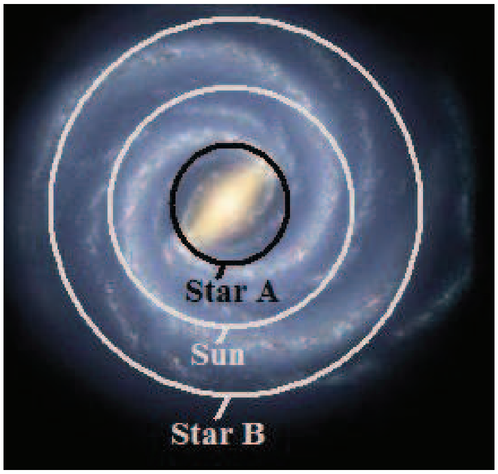
Part II: Motions of Stars

One way to estimate the amount of mass in a spiral galaxy is by looking at how much light the galaxy gives off. Where there is more light there must be more stars and hence more mass. When astronomers measure the amount of light coming from different parts of the galaxy, they find that stars are very concentrated in the central bulge of the galaxy and very spread out as you look outward through the disk.

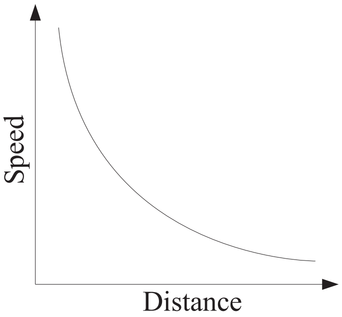
7) Based on the information provided above, where do you expect most the mass of a galaxy to be concentrated?

At right is a drawing of the Milky Way. The orbits of three stars are labeled. Star A is a star on the edge of the Milky Way's bulge. The Sun's orbit is shown at approximately the correct position. Star B is a star located farther out in the disk than the Sun.

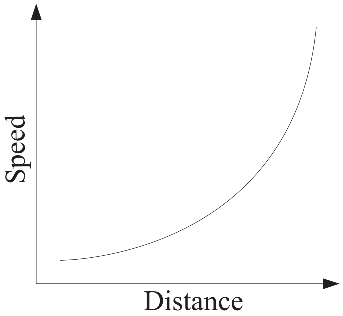
8) Based on your answer to Question 7 and the location of each star from the center of the galaxy, rank how you think the orbital speeds of Star A, Star B, and the Sun would compare from fastest to slowest. Explain your reasoning.



A graph of the orbital speed of stars versus their distance from the galaxy's center is called a *rotation curve*. Here are two possible rotation curves.



Rotation Curve 1

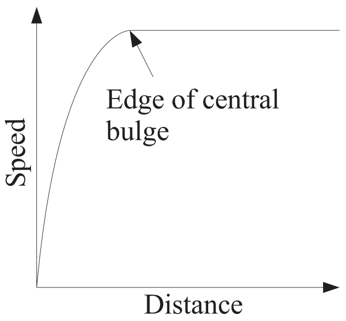


Rotation Curve 2

9) Which of the above rotation curves best represents the relationship you described in Question 8? Explain your reasoning.

Dark Matter

Astronomers were surprised when they saw the real rotation curve for the Milky Way Galaxy (MWG). The rotation curve at the right is more like the MWG's real rotation curve.



10) Based on their orbital distance from the center of the galaxy, make three dots on the above rotation curve to represent Star A, Star B, and the Sun. Be sure to label which mark belongs to each star.

11) Describe how the real rotation curve for the MWG is different from the rotation curve that you chose in Question 9.

12) Using the real rotation curve for the MWG, provide a new ranking for the orbital speed of Star A, Star B, and the Sun, from fastest to slowest. Describe any differences between this ranking and the one you provided in Question 8.

13) Based on your answers to question 12, would you say that the mass of the Milky Way Galaxy is concentrated at its center (as is the case with our Solar System)? Explain your reasoning.

14) Based on the MWG's real rotation curve and your answers to Questions 11-13, is the gravitational force felt by the MWG's stars greater than, less than, or about the same as what you expected? Explain your reasoning.

Dark Matter

15) Two students are debating their answers to the previous questions:

Student 1: *Stars far from the center of the Milky Way are all moving at about the same speed. If most of the Milky Way's mass was located in its center, then stars far away from the center would orbit slower than stars closer to the center. Since this is not what we see, this must mean there is more mass throughout the outer regions of the galaxy than we can see. This also means that the Milky Way's stars feel a greater gravitational force than we originally expected.*

Student 2: *I disagree. There are fewer stars in the outskirts of the Milky Way than in the center, so there's less mass out there than at the center. Most of the Milky Way's mass must be at its center. So, since the stars are all going about the same speed, where the mass is located must not affect their speed. The gravitational force these stars feel is exactly what astronomers expected.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

16) Astronomers initially thought there was more mass in the center of the galaxy than in the disk because there's more light coming from the galaxy's center than its disk. Describe how astronomers' observations show these initial ideas are wrong. Explain your reasoning.

17) Is there more or less mass in the Milky Way's disk and halo than we can see? Explain your reasoning.

Appendix I

Fall 2010 Scoring Rubrics

The scoring rubrics for the fall 2010 begin on the next page.

Form A

Item 1: Which graph or graphs show a universe that is expanding at a constant rate? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the expansion (or contraction) rate of the universe.
>
> **Correct answer:** F - This graph shows an expanding universe because galaxies that are farther away are moving away from us (positive velocity) faster than closer galaxies. Since the slope is constant, the rate at which the universe is expanding must be constant.

Item 2: Which graph or graphs show a universe that is contracting at a constant rate? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the expansion (or contraction) rate of the universe.
>
> **Correct answer:** B - This graph shows a contracting universe because galaxies that are farther away are moving toward us (negative velocity) faster than closer galaxies. Since the slope is constant, the rate at which the universe is contracting must be constant.

Item 3: Which graph or graphs show a universe that is expanding at a faster and faster rate over time? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the changing expansion (or contraction) rate of the universe, relative to lookback time.
>
> **Correct answer:** A - The rate at which the universe is expanding must be changing over time since the slope of the graph is changing with respect to distance. A universe expanding faster and faster over time must have a slope that gets steeper over time. This means the slope must be flatter at large distances since we are looking further back in time as we look farther away in distance.

Item 4: Which graph or graphs show a universe that is expanding at a slower and slower rate over time? Explain your reasons for making your selections.

> **What this item measures:** This item measures whether or not students can use Hubble plots to reason about the changing expansion (or contraction) rate of the universe, relative to lookback time.
>
> **Correct answer:** D - The rate at which the universe is expanding must be changing over time since the slope of the graph is changing with respect to distance. A universe expanding slower and slower over time must have a slope that gets flatter over time. This means the slope must be steeper at large distances since we are looking further back in time as we look farther away in distance.

The scoring rubric for these four items is on the next page.

Scoring rubric for items 1a-d.

Overall Score	**Answer Choice**	**Reasoning Categories**
3 - Student gives the correct answer with the correct reasoning.	A - Student selects graph A.	I - Description includes all or part of the following statement: "† is ± and is * at a ‡." *See below for a complete explanation of this reasoning category and symbols.*
	B - Student selects graph B.	
2 - Student gives the correct answer with incorrect or incomplete reasoning.	C - Student selects graph C.	II - Student talks about how when we look deep into space we are looking back in time.
	D - Student selects graph D.	III - Student says her/his reasoning is opposite his/her reasoning on the previous item.
1 - Student gives an incorrect answer.	E - Student selects graph E.	IV - Student says s/he is guessing and/or unsure of why s/he chose that answer.
	F - Student selects graph F.	
0 - Student gives no answer, offers irrelevant information, or claims s/he has no idea how to answer.	G - Student selects graph G.	V - Student says s/he can't give an answer without more information.
	H - Student selects graph H.	VI - Student gives no reasons for answer.
	Z - Student selects no graph.	VII - Student gives some other reason not captured by categories I-X above.

Each response receives an overall score (0-3) as when as a letter (A-H or Z) to denote which graph the student chose. A Roman numeral (I-VII) is also assigned to each response to mark the student's reasoning. Reasoning Category I is further subdivided, as described on the next page.

Reasoning Category I explained - Many students justify their selections by noting that the graph or some aspect of the graph is 1) positive or negative (i.e. above or below the x-axis), 2) increasing, decreasing, constant, or curved, and 3) changing with an increasing, decreasing, or constant rate. Some students include all of these elements in the reasons they wrote; most only have a subset of these reasons. To better distinguish between these various possibilities, there are various subclasses for Reasoning Category I:

When the student writes "† is ± and is * at a ‡," append, when appropriate, the following symbols to I.

† = v *if the student discusses velocity or speed*
= d *if the student discusses distance*
= s *if the student discusses slope*
= r *if the student discusses rate*
= ℓ *if the student discusses the line*
= g *if the student discusses the graph*
= o *if the student discusses some other feature*
= i *if the student discusses an unspecified feature or "it"*

± = + *if the subject † is positive (above the x-axis)*
= - *if the subject † is negative (below the x-axis)*

* = INC *if the subject † is increasing*
= DEC *if the subject † is decreasing*
= CON *if the subject † is constant/straight*
= EXP *if the subject † is curved or exponentiating*

‡ = R1 *if the rate at which subject † changes increases*
= R2 *if the rate at which subject † changes decreases*
= R3 *if the rate at which subject † changes stays constant*

Example responses and scores:

A response to item 1a: *"F. The line is straight, making the expansion constant and it is increasing both Velocity and Distance."*

> **How this response is scored:** 3F_I_lCON_vINC_dINC. This student gave the correct answer (F) and a correct reason. She thus receives an overall score of 3. In her answer, she discusses how the line (l) is constant (CON), the velocity (v) is increasing (INC), and the distance (d) is increasing (INC).

A response to item 1b: *"B, graph constantly going down."*

How this response is scored: 2B_I_gDEC_R3. This student gives the correct answer (B) but with an incomplete reason, thus leading to an overall score of 2. In her answer, she says the graph (g) is going down (DEC) constantly (R3).

A response to item 1c: *"D"*

How this response is scored: 1D_VI. This student gave the wrong answer (D, hence leading to an overall score of 1) and gave no reason (VI).

A response to item 1d: *"A,B,E. Velocity is decreasing as the distance is increasing"*

How this response is scored: 1_ABE_vDEC_dINC. This person earns an overall score of 1 because he gave the wrong answer (A, B, and E). His reason talks about the velocity (v) decreasing (DEC) and the distance (d) is increasing (INC).

Item 5: Use the blank graph provided below to draw what your think Figure 1 would look like if the universe had been expanding twice as fast **.** Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

> **What this item measures:** This item probes whether or not a student knows that the slope of a line on a Hubble plot should steepen if the expansion rate increases.

Scoring rubric for item 5.

Score	Response Includes	Reasoning Elements
2	- correct answer (steeper slope).	I - Student draws or discusses a steeper slope.
1	- incorrect answer or incorrect or incomplete reasons for choosing the correct answer.	II - Student draws or discusses a flatter slope. III - Student draws or discusses an unchanged slope. IV - Student draws a line with a non-constant slope.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	V - Student draws a line whose slope is ≤ 0. VI - Student draws a line whose y-intercept ≠ 0. VII - Student talks about increased velocity. VIII - Student talks about decreased velocity. IX - Student talks about velocity staying the same. X - Student talks about increased distance. XI - Student talks about decreased distance. XII - Student talks about distance staying the same. XIII - Student compares variable *x* to variable *y* by saying *x* is more than *y.* XIV - Student says distance and velocity change by the same amount. XV - Student says s/he can't answer the question without a time axis/variable. XVI - Student gives irrelevant information. XVII - Student gives some other reason not specified above. XVIII - Student has no idea. XIX - Answer field is blank.

Example responses and scores:

"The line would be much steeper b/c for every 1 unit of distance it expanded, the velocity would increase twice as much." [Student also draws a straight line with a steeper slope than Figure 1.]

> **How this response is scored:** 2_I_VII_XIII_X. This student draws a steeper line and says the line should be steeper (I). She also says both the distance (X) and velcocity (VII) increase,

although velocity increases more than distance (XIII).

"Same line, but units on graph will be twice as much" [Student also draws a straight line with the same slope as Figure 1.]

How this response is scored: 1_III_XIV. This student drew and discussed a slope that was the same as in Figure 1 (III). He also said that the values of both velocity and distance double (XIV).

Item 6: Use the blank graph provided below to draw what you think Figure 1 would look like for our universe if it took much longer to reach its current size. Explain the reasoning behind the graph you drew. If you don't have enough information to do this, explain what else you need to know.

What this item measures: This item probes whether a students knows that, in order for the universe to be older, it would need longer to reach its current size, which in turn means the expansion rate and slope of the Hubble plot both have to be smaller.

Scoring rubric for item 6.

Score	Response Includes	Reasoning Elements
3	- correct answer (flatter slope) - reason includes all of the following: 1) the universe took longer to reach its current size, 2) which means it had a slower expansion rate/velocity, and so 3) the graph has a flatter slope	I - Student draws or discusses a positive slope.
		II - Student draws or discusses a negative slope.
		III - Student draws or discusses a slope of zero.
		IV - Student draws or discusses a non-constant slope.
		V - Student draws or discusses a steeper slope.
		VI - Student draws or discusses a flatter slope.
		VII - Student draws or discusses an unchanged slope.
		VIII - Student draws a line whose y-intercept ≠ 0.
2	- correct answer (flatter slope) - incorrect or incomplete reason	IX - Student talks about a faster expansion rate.
		X - Student talks about a slower expansion rate.
		XI - Student talks about an unchanged expansion rate (i.e. expansion rate is the same as Figure 1).
1	- incorrect answer	XII - Student talks about an expansion rate that speeds up.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	XIII - Student talks about an expansion rate that slows down.
		XIV - Student talks about a constant expansion rate (i.e. expansion rate doesn't change over time).
		XV - Student talks about an increased velocity/speed.
		XVI - Student talks about a decreased velocity/speed.
		XVII - Student talks about a constant velocity/speed.
		XVIII - Student talks about an increased distance.
		XIX - Student talks about a decreased distance.
		XX - Student talks about a constant distance.
		XXI - Student talks about what happened to the universe in the past.
		XXII - Student talks about what will happen to the universe in the future.
		XXIII - Student says there's not enough information.
		XXIV - Students says the universe's age is irrelevant.
		XXV - Student talks about the time the universe needs to reach its current size.
		XXVI - Student says s/he needs to know how the expansion rate changes.
		XXVII - Student says there's no time variable/axis.
		XXVIII - Student says the significance of the universe's age is unclear.
		XXIX- Student gives irrelevant information.
		XXX - Student gives some other reason not specified above.
		XXXI - Student has no idea.
		XXXII - Answer field is blank.

Example responses and scores:

"smaller slope = longer time to expand due to slower expansion velocity" [Student also draws a straight line with a shallower slope than Figure 1.]

> **How this response is scored:** 3_I_VI_X. This student gets a 3 because she drew a line with a positive slope (I), the slope was flatter than the slope in Figure 1 (VI), and she said the expansion rate must have been slower (X).

"If it is taking twice as long, the slope needs to be 1/2 the original, so it needs to be more flat than the original." [Student also draws a straight line with a shallower slope than Figure 1.]

> **How this response is scored:** 2_I_VI. This student also drew and discussed a line with a positive slope (I) that is flatter than the slope of the line in Figure 1 (VI). However, this student did not mention the expansion rate of the universe, so he gets a 2.

"There is a lower velocity to distance rate" [Student also draws a graph like Graph B in items 1-4.]

> **How this response is scored:** 1_II_XVI. This student drew a line with a negative slope (II) and talked about a lower velocity (XVI).

Form B

Item 1: Explain, in as much detail as possible, what astronomers mean when they say "the universe is expanding." Provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes students' conceptions of how the universe increases in size. Does a student have an expert-like conception, in which case s/he talks about how galaxies are getting farther from each other due to the stretching (expansion) of space (or spacetime)? Does the student incorrectly view the expansion as affecting the distances between all objects, including stars, planets, and other objects smaller than galaxies? Does the student incorrectly think that the "expansion of the universe" refers to growth in our knowledge or the creation of new objects/locations over time?

The scoring rubric for item 1 is on the next page.

Scoring rubric for item 1.

Score	Response Characteristics	Reasoning Elements
4	Students in category 4 meet the same criteria as students in category 3 EXCEPT they only discuss increasing distances and/or redshifts in the context of galaxies (and not planets, stars, or other objects smaller than galaxies). Students may be placed in category 4 without explicitly saying that • the universe has no center, • the universe has no edge, • the Big Bang refers to the evolution of the universe from a hot, dense state. If the student does discuss one of these elements and makes an incorrect statement then s/he cannot be placed in category 4.	I - Student says the size of the universe increases over time. II - Student says the universe has a center. III - Student says the universe has no center. IV - Student says the universe has an edge. V - Student says the universe has no edge. VI - Student talks about redshifts/Doppler shifts. VII - Student says space(time) is growing/stretching. VIII - Student talks about the movement of galaxies and/or their increasing distances.
3	Students in this category say that • The universe increases in size and/or • The distances between all objects increase. Additionally, they must make at least one of the following claims and not contradict the others: • The universe has no center. • The universe has no edge. • The Big Bang refers to the evolution of the universe from a hot, dense state.	IX - Student talks about the movement of stars and/or their increasing distances. X - Student talks about the movement of planets and/or their increasing distances. XI - Student talks about the movement of objects (something unspecified or not a star, planet, or galaxy) and/or their increasing distances. XII - Student says the distances between everything increase.
2	Students in this category say that the universe increases in size and either provide no other information or claim one or more of the following: • The universe has a center. • The universe has an edge. • The Big Bang is an explosion. • The distances between all objects or objects smaller than galaxies increases.	XIII - Student says farther objects move away faster. XIV - Student talks about the Big Bang. XV - Student talks about an explosion. XVI - Student says the early universe was once hot, small, and/or dense. XVII - Student says we learn more about the universe over time.
1	Students in this category describe the expansion of the universe as something other than the universe getting larger over time. The two most popular answers that fit in this category are • Expansion refers to learning more about the universe over time. • Expansion refers to the creation of new objects over time.	XVIII - Student talks about how we are looking further back in time as we look farther into space. XIX - Student says new things are created in the universe over time. XX - Student gives irrelevant information. XXI- Student gives some other reason not specified above. XXII - Student has no idea.
0	Students in this category write nothing (the answer field is blank), or they provide information that doesn't answer the question, or they say they have no idea and provide no further information.	XXIII - Answer field is blank or the student provided no reason or explanation.

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response will contain every element in a given score category, you must also list the reasoning elements the student uses.

Example responses and scores:

"There was a man named 'hubble' who had a telescope and observed galaxies 'red shifting' or moving away. Now when he observed this he had the best telescope and plotted the rate of speed of expansion. He noticed that although the closer galaxies are red shifting at a higher rate, the most far away galaxies were moving away (Red shifting) at a much higher speed. From this observation astronomers deducted the universe is expanding from all far away galaxies are red shifting At a higher speed"

> **How this response is scored:** 4_VI_VIII_XIII. This student talked about redshifts (VI), how distances between galaxies are increasing (VIII), and how the farther away galaxies are, the faster they move away from us (XIII). Since the student talked about distances to galaxies increasing and did not make any incorrect statements about expansion, he receives an overall score of 4.

"'the universe is expanding' because as time goes on, matter is moving further and further away from other matter. Temperature and density have decreased over time as the universe is expanding and matter has moved away"

> **How this response is scored:** 3_XI_XVI. This student talks about otherwise unspecified pieces of matter moving away from one another (XI) and talks about how the universe used to be hotter and dense (XVI). While none of these pieces of information are necessarily wrong, she only gets an overall score of 3 because she does not specify whether expansion only affects the distances between galaxies or whether the distances between smaller objects are also affected.

"There was a big bang which exploded out everything in the universe. The leading edge of this bang has been expanding ever since."

> **How this response is scored:** 2_XIV_XV_XII_IV. This student talks about the Big Bang (XIV) as an explosion (XV). His answer indicates that distances between everything are increasing (XII) and he mentions an edge (IV). These reasoning elements give him an overall score of 2.

"I don't think that it is acculuy expanding in a physical sense, but instead or knowledge of the univers and the areas that we have discovered is expanding with an increase in technology and invesments in sciens"

> **How this response is scored:** 1_XVII. This student denies that the universe is physically growing. Instead, he says that expansion refers to our increase in knowledge over time (XVII), placing him in category 1.

Item 2: Explain, in as much detail as possible, what astronomers mean by the "Big Bang Theory." Provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether students think of the Big Bang as the beginning of the expansion of the universe, NOT an explosion of pre-existing matter into empty space, nor as the origin of something smaller than the universe (the Earth, Solar System, Galaxy, etc.).

The scoring rubric for item 2 is on the next page.

Scoring rubric for item 2.

Score	Response Characteristics	Reasoning Elements
3	Students in category 3 talk about the Big Bang using at least one of the following elements: • It's the beginning of the universe, space, and/or time. • It's when (some) of the elements were created. • It refers to how the universe was once hot, small, and dense. Additionally, they also include at least one of these elements (and do not contradict any of the others): • It was not an explosion. • It was the beginning of expansion. • Matter formed from energy. • It encompassed all of the universe.	I - Student says the Big Bang is the beginning of the universe/everything in universe. II - Student says the Big Bang is the beginning of expansion. III - Student says the Big Bang was the beginning of something smaller than the universe. IV - Student says the Big Bang is an event that happened to something smaller than the universe. V - Student says the Big Bang is the beginning of space. VI - Student says the Big Bang is the beginning of time. VII - Student talks about the creation/production of elements. VIII - Student says the Big Bang was an explosion. IX - Student says the Big Bang was not an explosion. X - Student says matter existed before the Big Bang. XI - Student says there was a dense piece of matter before the Big Bang. XII - Student talks about matter coming together before the Big Bang. XIII - Student says matter formed from energy. XIV - Student says the early universe was hot, dense, and/or small. XV - Student says the Big Bang was an event that happened in empty space. XVI - Student gives irrelevant information. XVII - Student gives some other reason not specified above. XVIII - Student says s/he has no idea. XIX - Answer field is blank or the student provided no reason.
2	Students in category 2 talk about the Big Bang using at least one of the following elements: • It's the beginning of the universe, space, and/or time. • It's when (some) of the elements were created. • It refers to how the universe was once hot, small, and dense. Students in this category either provide no further information or they use one of the following elements: • The Big Bang was an explosion. • Matter existed before the Big Bang. • The Big Bang was an event that happened in empty space. Additionally, any student whose ideas about the universe do not fit into categories 3 or 1 should be placed here.	
1	Students in category 1 talk about the Big Bang as the beginning of something smaller than the universe or and event that happened to something smaller than the universe.	
0	Students in category 0 write nothing (the answer field is blank), or they provide information that doesn't answer the question, or they say they have no idea and provide no further information.	

NB: You may assign multiple reasoning elements to a single response.

NB: The response characteristics are not meant to be an exhaustive list of everything contained within a students' response. Instead, they are meant to be guidelines to the common features of responses in each category. Because not every response with contain every element in a given score category, you

must also list the reasoning elements the student uses.

Example responses and scores:

"The Big Bang Theory, in a nutshell, says that the universe started out as very hot + dense, and over time that energy was converted into matter as the distance between matter increased and as the universe expanded."

> **How this response is scored:** 3_XIV_XIII_II. This student gets and overall score of 3 because in her response she mentioned expansion (II), the formation of matter from energy (XIII), and how the universe was once hot and dense (XIV).

"Big Bang Theory (not the TV show) means that all the stuff, matter, energy, whatever that exists in the universe now was once scrunched up into a teeny tiny little thing which then exploded and expanded rapidly."

> **How this response is scored:** 2_X_XI_II_VIII. This student receives a 2 for this response because she said the Big Bang is the beginning of expansion (II), a dense piece of matter existed before the Big Bang (X and XI), and the Big Bang was an explosion (VIII).

"'the big bang theory' is when an asteroid that was headed toward earth struck the earth and every thing that was alive died --- then as time went on things started growing and living again."

> **How this response is scored:** 1_IV. This response earns an overall score of 1 because it talks about the Big Bang as an event that happened to something smaller than the universe (IV).

"Big Bang theory means an idea or a way of critical thinking between astronomers, and/or scientists to the way they see the earth and universe through telescope or even without one through basic observation."

> **How this response is scored:** 0_XVI. This response does not really answer the question and gives a lot of irrelevant information (XVI).

Item 3: Each dot in the picture on the left is a galaxy. The Milky Way Galaxy (the one we live in) is at the center of the picture. All of the galaxies inside the circle can be seen from Earth. Any galaxies that exist outside the circle are so far away that their light has not had time to reach Earth. Describe what inhabitants of Galaxy X probably see when they look in the direction of the arrow.

> **What this item measures:** This item looks at whether students realize there are galaxies beyond our observable universe and that each location should have its own observable universe that looks like any other observable universe.

Scoring rubric for item 3.

Score	Response Includes	Reasoning Elements
2	- inhabitants of Galaxy X see something similar to what we see or they see more stars and/or galaxies	I - Student says they'd see more galaxies. II - Student says they'd see more stars. III - Student says they'd see more planets. IV - Student says they'd see more "objects" (otherwise unspecified) V - Student says they'd see something similar to what we see. VI - Student says they'd see things we cannot see. VII - Student says they'd see nothing/blackness/empty space. VIII - Student says they'd see objects separated by greater distances than what we see. IX - Students say Galaxy X's observable universe extends into regions outside of our observable universe. X - Student gives irrelevant information. XI - Student gives some other reason not specified above. XII - Student has no idea. XIII - Answer field is blank.
1	- does not specify if inhabitants of Galaxy X see something similar to what we see or they see more stars and/or galaxies - says the universe outside our observable universe is not similar to our observable universe	
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	

Example responses and scores:

"Galaxy X probably sees many galaxies in that direction. They probably look just as surrounded as Earth."

> **How this response is scored:** 2_I_V. This student says Galaxy X would see more galaxies (I) and their overall view would probably be similar to our view (V).

"They probably see nothing since there is no light able to reach there."

> **How this response is scored:** 1_VII. This student gets a score of 1 because she says inhabitants of Galaxy X would see nothing (VII).

Item 4: Circle the sentence that best describes the universe at the time of the Big Bang:

a) In the beginning, there was space in the universe surrounding the location of the Big Bang but this space was empty of all matter.

b) In the beginning, there was space in the universe surrounding the location of the Big Bang and matter already existed in this space.

c) I think of the Big Bang differently than a or b.

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes whether or not students think matter and space existed before the Big Bang.

Scoring rubric for item 4.

Score	Response Includes	Reasoning Elements
3	- correct answer (C) - says there was no space before the Big Bang - says there was no matter before the Big Bang	I - Student says space existed before the Big Bang. II - Student says there was no space before the Big Bang/no space outside of the Big Bang/all of space was part of the Big Bang.
2	- correct answer (C) - incorrect or incomplete reason	III - Student says space outside the Big Bang is necessary for the Big Bang. IV - Student says matter existed before the Big Bang.
1	- incorrect answer	V - Student says matter did not exist before the Big Bang.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	VI - Student says matter existing before the Big Bang is necessary for the Big Bang. VII - Student says matter cannot be created or destroyed. VIII - Student says all matter was once in one spot. IX - Student says expansion = matter filling empty space. X - Student says it is unlikely that matter existed anywhere else but the Big Bang. XI - Student says you can't make something out of nothing. XII - Student says time began with the Big Bang. XIII - Student says the Big Bang is the expansion of space. XIV - Student says the Big Bang is the beginning of the universe. XV - Student says the Big Bang was an explosion. XVI - Student says there was nothing before the Big Bang/everything began with the Big Bang. XVII - Student says we don't/can't know. XVIII - Student gives irrelevant information. XIX - Student gives some other reason not specified above. XX - Student has no idea. XXI - Answer field is blank.

Example responses and scores:

"C. At the instant of the big bang both space and matter came into existance in this universe. At the instant of the big bang there was nothing but the singularity. An instant after that space, time, and matter came into existance."

How this response is scored: 3_C_II_XII_V. This student earns a 3 because he chose the right answer (C) and defends his selection by saying there was no space (II), matter (V), or time (XII) before the Big Bang.

"C. infinitely small, infinitely dense speck of matter."

How this response is scored: 2_C_IV_VIII. This student chooses the right answer (C), but says there was matter before the Big Bang (IV) and it was all in one spot (VIII).

"B. Matter cannot be created or destroyed."

How this response is scored: 1_VII. This student circled a wrong answer (B) and said matter cannot be created or destroyed (VII).

Item 5: Independent of whether we know its true location, is there a center to the universe? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item examines whether or not students think the universe has a center and why.

Scoring rubric for item 5.

<table>
<tr><th>Score</th><th>Response Includes</th><th>Reasoning Elements</th></tr>
<tr><td>3</td><td>- rejects the idea of a center because the universe is infinite/has no edges
- rejects the idea of a center because the universe looks the same from all directions</td><td rowspan="4">I - Student says the universe is infinite/has no edges.
II - Student says the universe is the same everywhere.
III - Student says the Sun is at the center.
IV - Student says everything has a center/must have a center.
V - Student says the center is where the Big Bang happened/where the universe began/where the universe is expanding from.
VI - Student says things orbit the center of the universe.
VII - Student says our ignorance prevents us from seeing the center.
VIII - Student says technological limitations prevent us from seeing the center.
IX - Student says there is no center or the center changes because of expansion/objects are in motion.
X - Student says there is no center since the universe has no shape/an irregular shape/unknown shape.
XI - Student gives irrelevant information.
XII - Student gives some other reason not specified above.
XIII - Student has no idea.
XIV - Answer field is blank.</td></tr>
<tr><td>2</td><td>- rejects the idea of a center for a reason other than those listed in category 3</td></tr>
<tr><td>1</td><td>- says universe has a center or unclear if idea of center rejected
- says you can't see the center due to our ignorance/technological limitations
- says the center changes with expansion</td></tr>
<tr><td>0</td><td>- nothing (answer field left blank).
- information that does not answer the question.
- no idea.</td></tr>
</table>

Example responses and scores:

"no I think the universe is this object with no center or edge it is constantly going"

> **How this response is scored:** 3_I. This student gets a 3 because she says the universe has no edges (I).

"no. Because it's a 4D universe it doesn't have a known shape but it does not have a center because we don't know where exactly the location of the universe is."

> **How this response is scored:** 2_X. This student denies a center, but defends his answer by saying the universe has an unknown shape (X).

"yes, because the expansion had to start from a center point"

> **How this response is scored:** 1_V. This student says there is a center and that's where expansion started from (V).

Item 6: If you could travel to any location in the universe, could you go to a place where there would be no galaxies in front of you? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item probes whether students think there is an edge to the universe (in the sense that there is an end to the distribution of galaxies).

Scoring rubric for item 6.

Score	Response Includes:
3	- no, there will always be galaxies around you
2	- yes, if you go to a black hole or some other unintended or not quite correct reason.
1	- yes, there are regions where distribution of galaxies peters out/where there are no galaxies
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.

NB: There are no reasoning elements listed since each score is associated with its own unique reasoning element.

Example responses and scores:
"No, there are galaxies everywhere."

> **How this response is scored:** This response falls into category 3 since the student says there will always be galaxies around you.

"Um...I suppose you could go to a black hole and stare into it. Black holes are really super dense and no light escapes them so you wouldn't be able to see galaxies."

> **How this response is scored:** This student gets a score of 2 because she talks about looking into a black hole.

"Yes that would be toward the outside of the universe."

> **How this response is scored:** This response earns a score of 1 because the student says there are regions with no galaxies.

Item 7: Which of the following statements (a - d) are true? Circle all that apply.

In general, the expansion of the universe causes ________.
a) the distances between planets in the solar system to increase.
b) the distances between stars in the galaxy to increase.

c) the distances between galaxies in the universe to increase.
d) None of the above.
Explain your reasoning for your choice(s).

What this item measures: This item tests whether or not students realize that the expansion of the universe only affects distances between galaxies, not distances between stars within a galaxy or planets within a solar system.

Scoring rubric for item 7.

Score	Response Includes:
2	- idea that expansion only affects the distances between galaxies.
1	- idea that expansion affects distances other than those on galactic scales or that we cannot tell.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.

NB: There are no reasoning elements listed since each score is associated with its own unique reasoning element.

Example responses and scores:

"C. Stars and planets move toward and away from each other in the universe."

How this response is scored: This student gets an overall score of 2 because she says that expansion only affects the distances between galaxies.

"A, B. Planets + stars expand their orbits. I'm not sure if galaxies expand."

How this response is scored: This student gets an overall score of 1 because she claims that the distances other than those between galaxies increase due to expansion.

Form C

Item 1: Has the temperature of the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize the temperature of the universe has cooled over time.

The scoring rubric for item 1 is on the next page.

Scoring rubric for item 1.

Score	Response Includes	Reasoning Elements
3	- claim that universe has cooled over time. - reason based on the expansion of the universe.	I - Student says the temperature increased. II - Student says the temperature decreased. III - Student says the temperature stayed the same.
2	- claim that universe has cooled over time. - no reasons given or incomplete or incorrect reasons.	IV - Student says the temperature changed, but doesn't specify if it went up or down. V - Student talks about expansion and/or how the density of the universe changed.
1	- any claim other than the universe has cooled over time.	VI - Student talks about the birth/formation of stars/a star.
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	VII - Student talks about the birth/formation of planets/a planet. VIII - Student talks about the birth/formation of galaxies/a galaxy. IX - Student talks about the birth/formation of unspecified objects. X - Student talks about the death of stars. XI - Student talks about the death of planets. XII - Student talks about the death of galaxies. XIII - Student talks about the death of unspecified objects. XIV - Student talks about changes during the lives of stars/a star. XV - Student talks about changes during the lives of planets/a planet. XVI - Student talks about changes during the lives of galaxies/a galaxy. XVII - Student talks about changes during the lives of unspecified objects. XVIII - Student talks about how the universe is big. XIX - Student talks about competing effects canceling out. XX - Student gives irrelevant information. XXI - Student gives other reason not specified above. XXII - Student says s/he has no idea. XXIII - Response field left blank.

NB**:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"The temp. has gotten cooler. Again assuming no matter is created in the expansion process of the universe than while the universe has expanded it has become less dense and as you become less dense you cool off."

> **How this response is scored:** 3_II_V. This student gets a 3 because he says the temperature decreases (II) and he explains this in terms of the expansion of the universe (V).

"It could be getting colder as time goes on because as the universe expands there is more and more area for the sun and other giant stars to heat, and also the sun is continuously collapsing on itself (I think?) which might mean its getting smaller, just like other giant stars in the universe are which could also be making it cooler as time progresses."

> **How this response is scored:** 2_V_XVIII_XIV_II. This student correct says that the universe is cooling (II) and relates this to its expansion (V). However, she also relates the cooling to changes during the lives of stars (XIV) and mentions the big region (XVIII) stars have to heat. She thus receives an overall score of 2.

"It has to have changed b/c Galaxies are colliding, New stars are formed, and many stars die. The temperature can vary."

> **How this response is scored:** 1_IV_XVI_VI_X. This student says the temperature changed, but does not specify if he thinks it went up or down (IV) so he gets an overall score of 1. His answer also contains information on how galaxies change (XVI), the formation of stars (VI), and the death of stars.

Item 2: How does the <u>total amount of matter</u> in the universe right now compare to the total amount of matter in the universe at the very beginning of the universe (the moment just after the Big Bang)? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize that the total amount of matter is not conserved in the universe and that the early universe was pure energy.

Scoring rubric for item 2.

Score	Response Includes	Reasoning Elements
3	- claim that there is more matter in the universe now. - reason based on the idea that the early universe was pure energy and/or contained no matter until the temperature dropped.	I - Student says there was no matter/only energy in the beginning. II - Student talks about the temperature cooling and/or that matter formed from energy. III - Student says there is more matter now because the universe is expanding. IV - Student says the amount of matter increases as objects form and/or evolve. V - Student says the amount of matter decreases as objects form and/or evolve. VI - Student says the amount of matter increases as objects interact and/or die. VII - Student says the amount of matter decreases as objects interact and/or die. VIII- Student says the amount of matter does not change. IX - Student gives irrelevant information. X - Student gives other reason not specified above. XI - Student says s/he has no idea. XII - Response field left blank or no reason given.
2	- claim that there is more matter in the universe now. - no reasons given or incomplete or incorrect reasons.	
1	- any claim other than there is more matter in the universe now.	
0	- nothing (answer field left blank). - information that does not answer the question. - no idea.	

***NB*:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"There is tons more matter now. Then there was tons of energy, but now it has transformed into matter."

> **How this response is scored:** 3_I_II. This student says there used to be a lot of energy in the universe (I) but that some of it was subsequently transformed into matter (II), so she gets an overall score of 3.

"I think there is more because the universe is expanding and things are being created along w/expansion to 'fill it up'"

> **How this response is scored:** 2_III_IV. This student says the amount of matter in the universe has increased, but explains this in terms of the expansion of the universe (III) and new objects forming (IV). Therefore, she gets an overall score of 2.

"There is just as much matter now then then. It has just expanded."

> **How this response is scored:** 1_VIII. This student gets a 1 because he says the amount of matter hasn't changed (VIII).

Item 3: Has the density of matter in the universe changed over time, or has it always been about the same? Explain your reasoning and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether or not students realize that the density of matter has decreased over time due to expansion..

Scoring rubric for Item 3.

Score	Response Includes	Reasoning Elements
3	- claim that the density has decreased over time - reason based on the universe becoming bigger over time	I - Student says the density increases. II - Student says the density decreases. III - Student says the density is constant.
2	- claim that the the density has decreased over time - no reasons given or incomplete or incorrect reasons	IV - Student says the density changes, but doesn't specify how. V - Students says the amount of matter in the universe has changed.
1	- any claim other than the density has decreased over time	VI - Student says the amount of matter in the universe has not changed.
0	- nothing (answer field left blank) - information that does not answer the question - no idea	VII - Student says the size of the universe has changed/stuff spread out. VIII - Student says the size of the universe has not changed/stuff hasn't spread out. IX - Student talks about objects forming over time. X - Student talks about objects evolving over time. XI - Student talks about objects dying/breaking up. XII - Student talks about the effects of gravity. XIII - Student talks about matter changing forms. XIV - Student gives some other reason not specified above. XV - Student gives irrelevant information. XVI - Student says s/he has no idea. XVII - Answer field is blank or no response given.

***NB*:** You may assign multiple reasoning elements to a single response.

Example responses and scores:

"It was more dense @ the beginning, and is now less dense b/c the universe has expanded."

> **How this response is scored:** 3_II_VII. This student gets an overall score of 3 because she correctly said that the density has dropped (II) due to the expansion of the universe (VII).

"I think it's got less dense I mean there was none, then the Big Bang and now it formed."

> **How this response is scored:** 2_II_V_IX. This student gets an overall score of II because, although he said the density decreased (II), he also talked about how the amount of matter has changed (V) and objects forming (IX).

"It has changed, I couldn't say if its gotten larger or smaller but since D=M/V, and the mass has not changed, but I'm pretty sure V has, the D would be different."

> **How this response is scored:** 1_IV_VI_VII. This student says the density changes, but can't say how (IV), so she gets an overall score of 1. She also talks about how the amount of matter has not changed (VI), how the size of the universe has changed (VII).

Item 4: Galaxy X and Galaxy Y are part of an expanding universe. Inhabitants of Galaxy X see a star explode in Galaxy Y. They determine that the light from the explosion took 8 billion years to reach Galaxy X. How far apart were Galaxy X and Galaxy Y when the star exploded in Galaxy Y?

a) less than 8 billion light-years apart
b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart
d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether a student can relate distances and lookback times in an expanding universe.

Scoring rubric for item 4.

Score	Response Includes	Answer Choice	Reasoning Elements
3	- right answer (A) - correct reason	A - Student chooses answer A.	I - Student talks about the expanding universe/galaxies getting farther apart.
2	- right answer (A) - no reasons given or incomplete or incorrect reasons		II - Student talks about galaxies getting closer together.
		B - Student chooses answer B.	III - Student says light needs 8 billion years to travel 8 billion light-years.
1	- incorrect answer(s)		IV - Student says light-years are shorter than normal years.
0	- nothing (answer field left blank - information that does not answer the question - no idea	C - Student chooses answer C.	V - Student says light-years are longer than normal years.
		D - Student chooses answer D.	VI - Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.
		Z - Student makes no selection.	VII - Student says light may be slowed by something between the two galaxies.
			VIII - Student says we need more information on how galaxies moving apart.
			IX - Student says we need more information on how galaxies moving together.
			X- Student says 8 billion light-years is the minimum distance between Galaxies X and Y.
			XI - Student talks about time Galaxy X needs to see explosion.
			XII - Student says there's not enough information.
			XIII - Student gives irrelevant information.
			XIV - Student gives other reason not specified above.
			XV - Student says s/he has no idea.
			XVI - Response field left blank or no reason given.

Example responses and scores:

"A. The Galaxies are spreading apart so the light from the explosion had to catch up with the othe Galaxy which ws speeding away, so it took longer."

> **How this response is scored:** 3_A_I. The student chose the correct answer (A) and defended his answer by discussing the effects of expansion.

"A. it was in the galaxy so closer"

> **How this response is scored:** 2_A_XIV. This student selected the right answer (A) but gave a reason that does not match any of the common reasoning elements (XIV).

"B. The sight of the explosion travels @ the speed of light to Earth. 8-billion years @ the speed of light means 8-billion light years away."

> **How this response is scored:** 1_B_III. This student selected the wrong answer (B) and discussed how light would travel 8 billion light-years in 8 billion years.

Item 5: How far apart are Galaxy X and Galaxy Y when Galaxy X sees the star explode?

a) less than 8 billion light-years apart b) exactly 8 billion light-years apart
c) more than 8 billion light-years apart d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

> **What this item measures:** This item measures whether a student can relate distances and lookback times in an expanding universe.

Scoring rubric for item 5.

Score	Response Includes	Answer Choice	Reasoning Elements
3	- right answer (C) - correct reason	A - Student choose answer A.	I - Student talks about the expanding universe/galaxies getting farther apart.
2	- right answer (C) - no reasons given or incomplete or incorrect reasons		II - Student talks about galaxies getting closer together.
		B - Student chooses answer B.	III - Student says light needs 8 billion years to travel 8 billion light-years.
1	- incorrect answer(s)		IV - Student says light-years are shorter than normal years.
0	- nothing (answer field left blank - information that does not answer the question - no idea	C - Student chooses answer C.	V - Student says light-years are longer than normal years.
		D - Student chooses answer D.	VI - Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.
		Z - Student makes no selection.	VII - Student says light may be slowed by something between the two galaxies.
			VIII - Student says we need more information on how galaxies moving apart.
			IX - Student says we need more information on how galaxies moving together.
			X- Student says we need more information on the rate of expansion.
			XI - Student says answer depends on where, exactly, events happen in Galaxies X and Y.
			XII - Student says there's not enough information.
			XIII - Student gives irrelevant information.
			XIV - Student gives other reason not specified above.
			XV - Student says s/he has no idea.
			XVI - Response field left blank or no reason given.

Example responses and scores:

"C. The galaxies are drifting away and have been now for 8 B years."

> **How this response is scored:** 3_C_I. This student selected the correct answer (C) and explained his answer in terms of the expanding universe (I).

"C. The light you see is older. In the time light must travel, the light you are seeing might not be there

anymore"

How this response is scored: 2_C_XIV. The student chose the right answer (C) but for a reason that does not match any of the listed reasoning elements (XIV).

"D. We do not know if the galaxies are moving toward or away from each other. The fact that the universe is expanding means it is more likely that the galaxies are farther apart from each other now than earlier. But there is no information pertaining to the direction of the galaxies in relation to each other."

How this response is scored: 1_D_I_VIII_IX. This student talks about the universe expanding (I) but feels he needs more information on whether the galaxies are moving toward (IX) or away (VIII) from one another.

Item 6: The universe is 13 billion years old when inhabitants of Galaxy X see the star in Galaxy Y explode. When these inhabitants look at the exploding star, they see what Galaxy Y was like when the universe is what age?

a) less than 5 billion years old b) exactly 5 billion years old

c) more than 5 billion years old d) there is not enough information to tell

Explain your reasoning for your choice and provide a drawing if possible to help illustrate your thinking.

What this item measures: This item measures whether a student can relate distances and lookback times in an expanding universe.

Scoring rubric for item 6.

Score	Response Includes	Answer Choice	Reasoning Elements
3	- right answer (B) - correct reason	A - Student choose answer A.	I - Student talks about the expanding universe/galaxies getting farther apart.
2	- right answer (B) - no reasons given or incomplete or incorrect reasons		II - Student talks about galaxies getting closer together.
		B - Student chooses answer B.	III - Student says light needs 8 billion years to travel 8 billion light-years.
1	- incorrect answer(s)		IV - Student says light-years are shorter than normal years.
0	- nothing (answer field left blank) - information that does not answer the question - no idea	C - Student chooses answer C.	V - Student says light-years are longer than normal years.
		D - Student chooses answer D.	VI - Student says light-years are different than normal years, but doesn't know or doesn't specify if they think light-years are longer or shorter than normal years.
		Z - Student makes no selection.	VII - Student says light may be slowed by something between the two galaxies.
			VIII - Student says we need more information on how galaxies moving apart.
			IX - Student says we need more information on how galaxies moving together.
			X- Student reasons that 13-8=5/light traveled for 8 billion years.
			XI - Student talks about what Galaxy Y is like after the explosion.
			XII - Student talks about the star exploding more than 8 billion years ago.
			XIII - Student talks about how it takes time to see events happen in the universe/lookback time.
			XIV - Student says expansion affects age.
			XV - Student says answer depends on where, exactly, events happen in Galaxies X and Y.
			XVI - Student says there's not enough information.
			XVII - Student gives irrelevant information.
			XVIII - Student gives other reason not specified above.
			XIX - Student says s/he has no idea.
			XX - Response field left blank or no reason given.

Example responses and scores:

"B. 13-8=5"

> **How this response is scored:** 3_B_X. This student used subtraction (X) to arrive at the correct answer (B).

"B. Different life forms could believe anything really. They could think its been 20 billion years. By guessing that it has been 5 billion years exactly would seem most logical."

> **How this response is scored:** 2_B_XVIII. This student chose the right answer (B) but for a reason that does not correspond to any of the above reasoning elements (XVIII).

"A. It's more than 8 billion L.Y. away so they are looking more than 8 billion years in the past."

> **How this response is scored:** 1_A_XII. This student said the explosion happpened more than 8 billion years ago (XII) and thus chose the wrong answer (A).

Form D

Item 1: Which graph (1-6) best represents how **planets** orbit the Sun?

What this item measures: This item measures whether or not students can identify the correct rotation curve for a solar system (Graph 3).

Scoring rubric for item 1.

Score	Response
2	- student selects graph 3
1	- student selects any graph other than graph 3
0	- nothing (answer field left blank) - information that does not answer the question - no idea

Item 2: Which graph (1-6) best represents how **stars** orbit the center of the galaxy?

What this item measures: This item measures whether or not students can identify the correct rotation curve for a galaxy (Graph 2).

Scoring rubric for item 2.

Score	Response
2	- student selects graph 2
1	- student selects any graph other than graph 2
0	- nothing (answer field left blank) - information that does not answer the question - no idea

Item 3: Rank the speeds at which **Planets** A, B, and C orbit the Sun:

Ranking Order: Fastest speed 1________ 2________ 3________ Slowest

Or, all the planets orbit at approximately the same speed. _____ (indicate with a check mark)

Explain your reason for ranking this way:

What this item measures: This item measures whether students can use the rotation curve they selected in item 1 to correctly compare the speeds of the three planets.

Scoring rubric for item 3.

Score	Response Includes	Reasoning Elements
2	- student correctly relates the speeds of the planets based on the graph s/he chose	I - Student says most of the Solar System's mass is in its center.
1	- student incorrectly relates the speeds of the planets based on the graph s/he chose	II - Student says most of the Solar System's mass is not in its center.
0	- nothing - information that does not answer the question - no idea	III - Student says planets travel at the same speed because they all take the same amount of time to orbit.
		IV - Student says planets travel at the same speed regardless of distance (no further explanation).
		V - Student says closer (farther) planets move faster (slower); no further explanation given.
		VI - Student says closer (farther) planets move faster (slower) because they have a shorter (longer) distance to travel.
		VII - Student says closer (farther) planets move faster (slower) because they need less (more) time to orbit.
		VIII - Student says closer (farther) planets move slower (faster) because all planets take the same amount of time to orbit.
		IX - Student says closer (farther) planets move faster (slower) because they feel more (less) gravity.
		X - Student says closer (farther) planets move faster (slower) because they feel more (less) centripetal force.
		XI - Student references the effects of gravity, but her answer doesn't fit into the above categories.
		XII - Student gives irrelevant information.
		XIII - Student gives other reason not specified above.
		XIV - Student says s/he has no idea.
		XV - Response field left blank or no reason given.

Item 4: Rank the speeds at which **Stars** A, B, and C orbit the galaxy.

Ranking Order: Fastest speed 1________ 2________ 3________ Slowest

Or, all the stars orbit at approximately the same speed. _____ (indicate with a check mark)

Explain your reason for ranking this way:

What this item measures: This item measures whether students can use the rotation curve they selected in item 2 to correctly compare the speeds of the three stars.

Scoring rubric for item 4.

Score	Response Includes	Reasoning Elements
2	- student correctly relates the speeds of the planets based on the graph s/he chose	I - Student says stars orbiting the galaxy act like planets orbiting the Sun/answer is the same as item 3.
	- student incorrectly relates the speeds of the planets based on the graph s/he chose	II - Student says most of the galaxy's mass is in its center.
	- nothing - information that does not answer the question - no idea	III - Student says most of the galaxy's mass is not in its center.
		IV - Student says most of the galaxy's mass is in the spiral arms.
		V - Student says most of the galaxy's mass is not in the spiral arms.
		VI - Student says most of the galaxy's mass is located outside/at the edges of the galaxy.
		VII - Student says most of the galaxy's mass is not located outside/at the edges of the galaxy.
		VIII - Student says stars travel at the same speed because they all take the same amount of time to orbit.
		IX - Student says stars travel at the same speed regardless of distance because of dark matter.
		X - Student says stars travel at the same speed regardless of distance (no further explanation given).
		XI - Student says closer (farther) stars move faster (slower); no further explanation given.
		XII - Student says closer (farther) stars move faster (slower) because they have a shorter (longer) distance to travel.
		XIII - Student says closer (farther) stars move faster (slower) because they need less (more) time to orbit.
		XIV - Student says closer (farther) stars move slower (faster) because all stars take the same amount of time to orbit.
		XV - Student says closer (farther) stars move faster (slower) because they feel more (less) gravity.

	XVI - Student says the gravity of the galaxy is not concentrated in the galaxy's center due to dark matter. XVII - Student says closer (farther) stars move faster (slower) because they feel more (less) centripetal force. XVIII - Student says stars farther out in the galaxy are affected more by dark matter than stars closer to the center. XIX - Student references the effects of gravity, but her answer doesn't fit into the above categories. XX - Student gives irrelevant information. XXI - Student gives other reason not specified above. XXII - Student says s/he has no idea. XXIII - Response field left blank or no reason given.

Item 5: Based on your previous answers, how is matter distributed in **<u>solar systems</u>**? Pick the best answer from the following choices (a-c).

a) Most of the matter in the solar system is located in the Sun.
b) Most of the matter in the solar system is evenly distributed throughout the Sun and planets.
c) My thinking is different than a and b.
Explain your reasoning.

What this item measures: This looks at whether or not the student understands where matter is concentrated in solar systems and why.

Scoring rubric for item 5.

Score	Response Includes	Reasoning Elements
3	- student chooses the correct answer (A) and her answer is consistent with her answer to item 1	I - Student says the Sun is big/the most massive object in the Solar System.
2	- student chooses the correct answer (A) and her answer is inconsistent with her answer to item 1	II - Student says the Sun has a large gravitational force.
1	- student chooses an incorrect answer	III - Student says the Solar System's gravity/mass is concentrated at the location around which everything orbit.
0	- nothing - information that does not answer the question - no idea	IV - Student says other non-Sun objects in the Solar System contribute to the mass of the Solar System.
		V - Student says the mass of all the non-Sun objects equals or exceeds the Sun's mass.
		VI - Student says most of the mass of the Solar System is at its edge.
		VII - Student gives irrelevant information.
		VIII - Student gives other reason not specified above.
		IX - Student says s/he has no idea.
		X - Response field left blank or no reason given.

Item 6: Based on your previous answers, how is matter distributed in **<u>spiral galaxies</u>**? Pick the best answer from the following choices (a-c).

a) Most of the matter in the galaxy is located in the center.
b) Most of the matter in the galaxy is located in the spiral arms.
c) My thinking is different than a and b.
Explain your reasoning.

What this item measures: This looks at whether or not the student understands where matter is concentrated in spiral galaxies and why.

Scoring rubric for item 6.

Score	Response Includes	Reasoning Elements
3	- student chooses the correct answer (C) and her answer is consistent with her answer to item 2	I - Student says the distribution of mass in the galaxy is like the distribution of mass in the Solar System/ the answer is the same as item 5.
2	- student chooses the correct answer (C) and her answer is inconsistent with her answer to item 2	II - Student says most of the mass is at the center because that's where the density is highest/galaxy is brightest.
1	- student chooses an incorrect answer	III - Student says there is little matter in the center.
0	- nothing - information that does not answer the question - no idea	IV - Student says most of the matter in the galaxy is at its edge/outskirts.
		V - Student says most of the matter in the galaxy is in its stars.
		VI - Student says matter is evenly distributed throughout the galaxy.
		VII - Student says most of the mass of the galaxy is dark matter.
		VIII - Student says there is more dark matter in the spiral arms/disk than anywhere else.
		IX - Student says dark matter is located outside the parts of the galaxy you can see.
		X - Student says matter is brought closer together as it moves toward the center.
		XI - Student talks about the black hole in the galaxy's center.
		XII - Student says most of the galaxy's gravity is located in its spiral arms.
		XIII - Student says gravity is evenly distributed throughout the galaxy.
		XIV - Student says the galaxy's gravity/mass is concentrated at the location around which everything orbits.
		XV - Student references dark matter, but her answer doesn't fit into any of the above categories.
		XVI - Student gives irrelevant information.
		XVII - Student gives other reason not specified above.
		XVIII - Student says s/he has no idea.
		XIX - Response field left blank or no reason

	given.

Item 7: Based on your answers for Questions 1-7, do **stars** orbiting the center of a galaxy act like **planets** orbiting the Sun? If yes, explain why. If no, explain why not.

What this item measures: This question looks at whether students can use their previous answers to compare the orbits of planets and stars and explain why galaxies, unlike solar systems, have flat rotation curves.

Scoring rubric for item 7.

Score	Response Includes	Reasoning Elements
3	- student answers no - student says that matter is distributed differently in solar systems and galaxies AND/OR that the velocities of stars/rotation curves of galaxies are different than the velocities of planets/rotation curves of solar systems AND/OR that her previous responses were different for solar systems and galaxies (and were also scored as correct)	I - Student says that orbiting stars act like orbiting planets. II - Student says that orbiting stars do not act like orbiting planets. III - Student says that orbiting stars do not act like orbiting planets, but only because of superficial reasons. IV - Student says stars move slower than planets.
2	- student says no - student's reasoning is incorrect or incomplete	V - Student says stars have irregular orbits/don't orbit a specific point, unlike planets.
1	- student says they are the same or similar	VI - Student says the distribution of matter in galaxies is different from the distribution of matter in solar systems.
0	- nothing (answer field left blank) - information that does not answer the question - no idea	VII - Student says the velocities of stars/rotation curves of galaxies are different than the velocities of planets/rotation curves of solar systems. VIII - Student talks about the black hole at the center of the galaxy. IX - Student references the effects of gravity, but her response does not fit into any of the previous categories. X - Student references the effects of dark matter, but her response does not fit into any of the previous categories. XI - Student says her previous responses were the same for the solar system and the galaxy. XII - Student says her previous responses were not the same for the solar system and the galaxy. XIII - Student gives irrelevant information. XIV - Student gives other reason not specified above. XV - Student says s/he has no idea. XVI - Response field left blank or no reason given.

Example responses and scores:

NB: All of these responses are taken from the same student.

Item 1: *"5"*

How this response is scored: 1. This student did not choose graph 3.

Item 2: *"2"*

How this response is scored: 2. This student chose graph 2.

Item 3: *"A, B, C. Most of mass of Solar System is located at the center of the Sun"*

How this response is scored: 2_I. This response is consistent with his response to item 1. He specified where most of the mass is located (I).

Item 4: *"most of the mass is not located in the center of the universe"*

How this response is scored: 2_III. This response is consistent with his response to item 2 and he said most of the mass is not in the center (III). I ignored the fact that he said "universe" and not "galaxy."

Item 5: *"A"*

How this response is scored: 3_X. A is the correct answer and it is consistent with his graph choice in item 1. He did not provide any reason for his answer (X).

Item 6: *"C. distributed throughout"*

How this response is scored: 3_VI. C is the correct answer and it is consistent with his graph choice in item 2. He said matter is distributed throughout the galaxy. (VI).

Item 7: *"NO, planets at farther distances orbit more slowly"*

How this response is scored: 3_II_VII. This student correctly says that the orbits of planets are different from the orbits of stars (II) because planets move slower at greater orbital distances (VII).