SIMULATING SCIENCE: USING A COMPUTER TO FACILITATE CONCEPTUAL CHANGE FOR MOON PHASES

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

Of the Requirements for the Degree

Doctor of Philosophy

By

William Robert Alexander, B.S. & M.S.

August 2014

Department of Curriculum, Instruction, and Special Education
Curry School of Education
University of Virginia
Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

This dissertation, Simulating Science: Using a Computer to Facilitate Conceptual Change for Moon Phases, has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Jennifer Chiu, Advisor

Randy L. Bell, Committee Member

Edward M. Murphy, Committee Member

Walter F. Heinecke, Committee Member

Joe Garofalo, Committee Member

Date

ABSTRACT

The teaching of lunar phases to pre-service elementary teachers is important since lunar phases and the relationship of the earth-moon-sun system are specifically stated in the national science education reform documents. Moreover, *Benchmarks for Science Literacy (Benchmarks)* (American Association for the Advancement of Science, 1993), *National Science Education Content Standards (NSES)* (National Research Council, 1996), as well as *Next Generation Science Standards (NGSS)* (National Research Council, 2013) give specific recommendations for making use of physical models to help students understand the geometry of the sun-earth-moon system that is critical for understanding phases of the moon. Each of these documents view concepts of lunar phases as foundational for understanding the sun-earth-moon system. Additionally, these national reform documents have been modified by many states to serve as a basis for their state science education standards.

Pre-service elementary teachers' conceptions of shape, sequence, and scientific understanding of the cause of moon phases improve both statistically and substantively after inquiry-based, conceptual change instructional intervention based on a computer simulation. More specifically, planetarium software was found useful to facilitate a constructivist approach to science education where the students constructed their own understanding of lunar phases by the use of this inquiry-based, conceptual change approach. Gains in conceptual understanding were measured by a semi-structured interview protocol as well as the LPCI and an Abridged-LPCI that is a version of the

LPCI that includes only the questions that align directly with the interview protocol. The Abridged-LPCI results correlate better than the LPCI with the semi-structured interview protocol. Although there were statistically significant correlations, they were not practically significant. Student interviews, as well as student daily written reflections, provided data that helped to illuminate the conceptual change process and are not available from forced-choice instruments.

Additionally, there are no statistical or substantive differences in achievement between students who collect moon observation data in a whole class setting versus those who collect this data working with the computer simulation in pairs. In general, the factors measured from the writing of student reflections were not significantly correlated with measures of conceptual gains although they can provide insights into the learning process.

And finally, the Styrofoam™ ball (scientific model) activity was found to be a significant instructional component that positively impacted improvement of scientific conceptions of moon phases. Only 2 of 12 possible correlations between student reflection and LPCI were statistically significantly positively correlated and none of the 12 possible correlations for the Abridged-LPCI were statistically significantly correlated with student reflections, although they were not practically significant. Moreover, the students generally report that student reflections are not an important factor for the instructional intervention to students' attainment of scientific conceptions of moon phases. Additionally, there were three new alternative lunar phase alternative conceptions that were identified in pre-instruction interviews. These new alternative conceptions include: Alternative Orbital Speed (the moon orbits the Earth in 24 hours, thus causing the lunar phases), Alternative No Orbit (the moon is in the same position relative to the Earth and Sun for all of the lunar phases), and Alternative Large Sun (the Sun is larger than the Earth, so light goes around the Earth causing the full moon). It was found that the semi-structured interview protocol was more able to measure a variety of alternative conceptions relating to lunar phases concepts when compared to a forced-choice instrument such as the LPCI or the Abridged-LPCI.

The above findings have implications to both practicing and pre-service teachers as well as allowing for several new areas of research related to lunar phases concepts and the elements related to teaching them.

DEDICATION

I would like to dedicate this study to my children: Ian, Jada, and Morgan. You have brought immense joy and meaning to my life.

ACKNOWLEDGEMENTS

I would like to thank all members of my committee for their valuable feedback and encouragement during this study. I appreciate the assistance in performing my statistical analysis that I received from Kane Nashimoto at James Madison University. A special thank you goes to Randy Bell for his patience and guidance during my time at the Curry School. Lastly, I would like to thank Geary Albright for opening up his classroom for this study and thus making my research possible.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Introduction

A lively debate has begun among college-level astronomy educators. Should phases of the moon even be included within a college-level astronomy course? Should grants be issued to educators to assess the teaching of lunar phases to college students? At odds are the "astronomy educators" and the "astronomers." The astronomy educator can be called the "traditionalist" and supports the notion that the basics of astronomy are important and there is a certain astronomical knowledge base that every educated individual should have. The astronomers (who also teach astronomy at the college level) can be viewed as the "modernists" and support the notion that modern astronomical discoveries promote a sense of wonderment that motivate students to study astronomy. This debate has been documented (Pasachoff, 2002). It is beyond the scope of the present work to fully address this debate, but it sets the context for the present investigation.

No such debate exists among K-12 educators and those that prepare these educators, since lunar phases and the relationship of the earth-moon-sun system are specifically stated in the national science education reform documents. *Benchmarks for Science Literacy (Benchmarks)* (American Association for the Advancement of Science, 1993), *National Science Education Content Standards (NSES)* (National Research Council, 1996), as well as *Next Generation Science Standards (NGSS)* (National Research Council, 2013) give specific recommendations for making use of physical models to help students understand the geometry of the sun-earth-moon system that is critical for understanding phases of the moon. Each of these documents view concepts of lunar phases as foundational for understanding the sun-earth-moon system. These national reform documents have been modified by many states to serve as a basis for their state science education standards.

The term “phases of the moon” can also be referred to as “lunar phases” and indeed the two terms are used interchangeably in the literature. The lowest order conception of lunar phases is the shapes that the moon appears to make as the lunar cycle progresses. The next conceptual level is the lunar orbit around the earth and how it relates to the lunar phases cycle. A slightly higher conceptual level is the relative positions of earth-moon-sun system and how it relates to the lunar phases. This level is generally accepted as the “reason” for the lunar phases as seen from earth. Related to the relative positions of the earth, moon and sun are ocean tides. This concept has been studied by only a handful of the investigations. The next conceptual level would be the relatively rare occurrence of a solar and lunar eclipse, which can be explained by the moon’s five-degree orbital inclination with respect to the earth’s orbit around the sun. The highest level of lunar phase teaching covered in the literature is the position of the moon in the sky for a particular lunar phase at a particular time of day. This level is seldom addressed in the literature. It is apparent from the literature that research into the teaching of phases of the moon is important to a significant number of science educators.

A common theme in the literature is instrument development to document alternative conceptions and to investigate various teaching approaches associated with the phases of the moon. The three main instruments related to determining students’ conceptual understanding of lunar phases included the Astronomy Diagnostic Test (ADT) developed by Huffnagel (2002), the Lunar Phase Concept Inventory (LPCI) developed by Lindell (2001), and a semi-structured interview protocol developed in Trundle, Atwood, and Christopher (2002) that specifically measures students' conceptual understanding of lunar phases. The ADT is an instrument designed to measure general astronomy knowledge of college students and is not specific to lunar phases. The ADT and LPCI are forced-choice instruments that take 30 minutes to administer. The literature has relied heavily on these instruments for empirical data to assess teaching effectiveness as well as determining frequency of alternative conceptions, although the two instruments have never been systematically compared. The reflections that students keep during instruction could provide a wealth of data and have not been studied to see if they can be used to assess or predict student outcomes.

Researchers have uncovered common alternative conceptions that are found among elementary as well as college students (Schoon, 1995). It is apparent that college students generally do not hold a scientific view as to the causes of the phases of the moon when they begin an introductory astronomy course (Bell & Trundle, 2006, 2007, 2008; Bisard, Aron, Francek, & Nelson, 1994; Kuethe, 1963; Lindell, 2003; Lindell & Sommer, 2003; Parker & Heywood, 1998; Schoon, 1992, 1995; Schoon & Boone, 1998; Trumper, 2000; Trundle et al., 2002, 2006, 2007a,b). The most prevalent alternative conception is that the lunar phases are caused by the Earth's shadow being cast on the moon (Trundle et al., 2002). Also, it is clear from the literature that students do not enter the college-level astronomy courses with as much knowledge about lunar phases as one might think (Trundle et al., 2002).

Several studies compared teaching approaches. Many of the studies considered non-content factors such as engagement of students (Barnett & Moran, 2002) or the thoughtfulness of student reflections (Abell et al., 2002). However Bell and Trundle (2006, 2007, 2008), Trundle et al. (2002, 2006, 2007a,b), and Trundle and Bell (2010) have shown that inquiry-based, conceptual change instruction that addresses the student's alternative conceptions can be an effective approach to teaching lunar phase concepts. Keys and Bryan (2001) describe inquiry-based, conceptual change science instruction as including: students posing questions, designing and carrying out investigations, analyzing their own data, developing a model to interpret their data, and lastly communicating their findings. Addressing students' alternative conceptions in order to replace them with the scientific explanations is known as the conceptual change approach to science teaching. Moreover, these studies provide details of their studies that used a standardized interview protocol instrument to measure gains in conceptual understanding of lunar phases. This makes it possible to compare various teaching approaches. To this end, Trundle and Bell (2005) developed a novel inquiry-based, conceptual change teaching technique using personal computers operating "planetarium software" that simulates the night sky and has shown great promise.

Personal computers can provide the opportunity to simulate difficult to observe phenomena in nature, such as extended observations of lunar phases. There is much research related to computers in the classroom, subsequently there is also a significant body of literature related to the use of computer simulations in the science classroom. Studies related to computer simulations of natural phenomena have been conducted at various grade levels and approaches: from the elementary to the college classroom, for specific content goals, and for specific pedagogical goals as well. In a recent review article, Smetana and Bell (2012) found that computer simulations can be as effective, and often more effective, than more traditional instructional methods to promote conceptual change.

Of particular importance are computer simulations used in teaching Space sciences that show the most improvement in conceptual understanding. "Space Science" is defined as primarily the study of our solar system, with a special emphasis on the planets, sun, and moons. A computer simulation will refer to any use of a computer to simulate a natural phenomenon from either earth or space science. These computer simulations will make use of a visual representation of the phenomenon which may, or may not be easily observable in nature. These simulations will attempt to give an accurate visual representation of the phenomenon within the limits of the available technology.

The use of computer-based simulation software in order to reproduce observations and/or simulations of space science events has been shown to be generally effective. The Virtual Solar System (VSS) project, in which virtual reality modeling software was used to model the solar system, was found to be an effective approach to teaching solar system concepts (Barnett, Barab, & Hay, 2002). The VSS used a constructivist approach to teaching solar system concepts. The constructivist approach, as used in science education, can be defined as the realization that scientific knowledge is built from a person actively thinking and that knowledge is not passively transferred from instructor to student (Starver, 1998). The social interaction between learners as well as instructor is central to this building of knowledge. The commercially available software *Starry Night©* has also been found to be quite effective in teaching space science topics, namely the phases of the moon. Additionally, there are initial indications that the use of *Starry Night©* can result in more scientific conceptual understanding of lunar phases compared to real sky observations (Bell & Trundle, 2006) and can additionally promote a constructivist-based approach to science education where the students can construct their own understanding of scientific phenomena. However, there has been no investigation into whether students learning lunar phases should be working in small groups, or even if perhaps an entire class could make their observations from the same computer used in demonstration mode. Chang (2003) showed that in a high school earth science class, the teacher-directed computer simulations were more effective than the student-controlled in earth science. However, Smetana and Bell (in press) showed that there was no significant difference between the two types of groups for a high school chemistry class. This could be an issue because the cost of multiple-computer site licenses may be too expensive for some school systems, while a single copy of the software could be relatively inexpensive.

Questions also remain regarding the duration of daily observations required for lunar phase instruction, regardless of using *Starry Night©* or making natural observations. The time required for observations has ranged from a high end of nine weeks (Trundle et al., 2002) in order to cover two full lunar cycles, to a low end of sixteen days (Bell et al., 2007) that ranged from two days before full moon to one day before new moon. Bell et al. (2007) showed that the use of *Starry Night©* could result in very favorable conceptual gains with only using the sixteen days of observations. In previous studies, the reasoning for the number of observations taken has not been well-defined. The reference *Physics by Inquiry* (McDermott, 1996) is often used as a basis for making observations of the sky, but in the work the only mention of duration of observations is "for the next few weeks, make the observations." (McDermott, 1996, p. 337) There is no specific rationale given for the duration of the observations, therefore the duration should be revisited. The number of observations should be based on how the observations will ultimately be used in the two dimensional (2-D) drawings as well as the three dimensional (3-D) manipulation of models that follows the observations since this is the critical portion of instruction where the students are constructing mental models of the earth-moon-sun system.

Statement of the Problem

Since the use of planetarium software such as *Starry Night*© can be cost-prohibitive for an entire classroom, there should be an investigation that studies the use of the software in a demonstration mode. Thus, fewer copies of the software would potentially need to be purchased if the demonstration modes performs as well as or even better than small groups students working at individual computers determined by measuring conceptual gains before and after instruction. Additionally, the use of the whole-class demonstration approach can facilitate learning and discussion which can lead to formative assessment of students so that the teacher can adapt discussion during instruction to best meet the student's needs (Smetana & Bell, 2012).

In order to address the lack of empirical data used to base the duration of lunar observations upon, an investigation should be made that will have students making observations from new to last (third) quarter moon to determine if the modeling of these observations using 3-D models will prepare students to correctly predict the shape, sequence and timing of the phases from last quarter to new moon.

Additionally, another area that has not been explored in previous research is to determine the most important factors of instruction. For instance, what role do student drawings made during instruction play in student conceptual understanding gains measured after instruction. Students should also be interviewed after instruction and asked what they felt the most important factors of instruction were. Investigating these issues will help to fill gaps in knowledge related to lunar phase teaching since the literature mostly includes data taken with instruments before and after instruction, not during instruction. Not much is known about how students construct their conceptual understanding during instruction so these topics merit further investigation.

The literature shows that the whole-class versus student-controlled computer simulation issue is not yet resolved, so there is a need for research to investigate conceptual gains of students when both approaches are compared. Additionally, both approaches should use inquiry based, conceptual change instruction.

The study will specifically address the previously mentioned deficiencies in the literature by answering the following research questions:

1. How do pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases change after inquiry-based, conceptual change instructional intervention based on a computer simulation?
2. Are there differences in achievement between pre-service elementary teachers who collect moon observation data in a whole-class setting versus those who collect this data working in pairs with the computer simulation?

3. How do pre-service elementary teachers perceive the relative value of each component of the instructional intervention as related to their attainment of scientific conceptions of moon phases?
4. How does the forced-choice LPCI instrument compare with the semi-structured Trundle protocol in measuring the conceptual understanding of lunar phases among pre-service elementary teachers?

Significance of Study

The argument for the importance of lunar phase teaching in general can be approached from a variety of angles. One might argue that if teachers need weeks or even months to address teach scientific conceptions of the moon, there will not be time to cover other topics such as black holes, sun quakes and starburst galaxies. On the other hand, phases of the moon can be considered as "foundational" and make a convenient gateway to outer space. Investigating approaches that could reduce the total number of class hours can help to provide more time to cover other topics in Space science. Perhaps it is not necessary to have students gathering data over two complete lunar cycles in order for them to develop a scientific understanding as to the cause of lunar phases. Maybe only the waxing phases for observations could achieve this reduction in classroom hours, while still hopefully providing a sufficient level of conceptual gains. Since lunar phases are directly addressed in the *Benchmarks* and the *NSES*, this study can have a direct impact on the teaching and learning of core science content.

The LPCI has never been used to purposefully sample students for more extensive interviews, so this is a novel approach. This is important since it is not practical to interview each participant before and after instruction. The LPCI will be an important tool to determine a representative sampling of higher, middle, and lower performing students, from pre-instruction results, that will be further interviewed. The pre-instruction LPCI will be performed before the pre-instruction interviews.

Researching effective approaches, such as whole-class demo vs. students working in small groups, can assist teachers with planning the number of computers that they need in their classrooms. This is important because a site license for *Starry Night*© to equip an entire computer lab could be quite expensive, while a single copy of *Starry Night*© can be purchased at a moderate cost for use with a lecture computer. If whole-class demos are effective, it may mean that teachers who lack the resources of multiple computers can make effective use of *Starry Night*© software to teach lunar phases. This might be particularly useful for practicing teachers at elementary schools with limited funding.

It will be important to find what factors are the most important in the learning process for lunar phases in order to inform future instruction. This will include understanding the relationship between student drawings made during instruction and conceptual gains. If there is indeed a direct relationship between the two, there would be an incentive for teachers to actively evaluate drawings as they are being made so as to direct the students in their inquiry process. Moreover, the teacher can better react to alternative conceptions of lunar phases that present themselves in student drawings during instruction. Additionally, the post-instruction interview can provide insight into the development of conceptual models related to lunar phases and how they change when students use the models to make predictions. Previous studies have not studied the development of mental models of lunar phases and this study provides a unique opportunity to do so.

Lastly, it will also be important to compare the LPCI with the Trundle semi-structured interview protocol because the LPCI takes less time to administer. To the extent that the LPCI could potentially replace the interview protocol, future studies could be expanded to include more participants.

CHAPTER 2: REVIEW OF LITERATURE

Introduction to the Literature

The body of literature concerning the teaching and assessment of lunar phases to college-aged students is diverse. The two types of groups studied college students that were taking astronomy to fulfill a general science requirement and also students enrolled in a course for teacher preparation in the sciences. Not only is there a major division within college-aged student populations that were studied, but there are also different paradigms which oversee the two major approaches to assessment. Those studying the general college students generally used quantitative methods that used paper and pencil instruments, where those studying pre-service teachers used a mixed approach of qualitative interviews that were generally coded and converted to quantitative data.

A particular challenge to performing a comprehensive literature review related to teaching lunar phase concepts has been the wide variety of journals in which research has been published. There were certain clusters of researchers who cited each others' work, but there were also examples of small research groups that only cited their own work. An example of an isolated research group would be the works by Zeilik and Hufnagel that were published in *Publications of the Astronomical Society of Australia*, *The American Journal of Physics*, and the online journal *Astronomy Education Review*. No citations were given by authors Zeilik and Hufnagel and their fellow researchers from journals such as the *Journal of Research in Science Teaching* and *Journal of Science Teacher*

Education. These journals are geared toward science teacher preparation and k-12 education and researchers such as Trundle and Stahly published in these. The lack of citations from Zeilik and Hufnagel could be because their work was geared towards the general college student, not the preservice teacher. Zeilik and Hufnagel also do not cite works from field of cognitive psychology by researchers such as Vosniadou and Brewer that are published in journals such as *Cognitive Science* and *Cognitive Development*. Likewise, researchers Vosniadou and Brewer do not cite works by Zeilik, Hufnagel, Trundle, or Stahly. Only the researchers dealing with preparation of pre-service teachers, such as Trundle and Bell, tend to make comprehensive literature reviews that encompass a variety of journals. For this study, the articles reviewed came from a variety of sources, however only empirical studies were reviewed. In addition to peer reviewed journal mentioned, sources included published dissertations, conferences proceedings (such as the National Association for Research in Science Teaching), and online peer-reviewed journals such as the *Astronomy Education Review*.

There are two distinct groups of college students that take introductory astronomy courses. The courses for pre-service teachers are unique in that students have much in common as far as their ultimate use of what they are learning. For most of these pre-service teachers, phases of the moon will be part of their responsibilities as teachers. The general student can be a much more diverse group. Typically “astro-101 courses” are populated with non-science majors taking astronomy for a science credit. But these general courses may also include science majors who may potentially choose astronomy as a profession. One might argue that the pre-conceptual understanding for these general students may vary a great deal when compared to the pre-service teachers. Studies such as that have not been conducted, but this likelihood makes it appropriate to separate the two types of studies.

The two different methodologies used in these studies are a result of the two major paradigms in science education research, which are qualitative and quantitative methodologies. A majority of the studies conducted with pre-service teachers used qualitative approaches. These methodologies included interviews, reflections and journals. To a lesser extent, quantitative methods have been used in the study of how pre-service teachers learn phases of the moon. Only quantitative methods have been employed to study general classes of astro-101. This is likely due to the large portion of these classes have large enrollments (N>100). The one theme that ties all these studies together is the determination of pre-conceptions relating to phases of the moon.

There is an emphasis on instrument development and documentation of lunar phase alternative conceptions and preconceptions in the literature. There are far fewer studies associated with approaches to teaching lunar phase concepts. Particular attention will be paid to the development of quantitative assessment instruments as well as qualitative research protocols.

As shown in Table 1, the approach to developing assessment instruments, as well as documenting lunar phase misconceptions, has been much more quantitative than qualitative. The term “mixed”, as indicated in Table 1, represents methods that qualitative methods, such as interviews, that have been coded and converted into quantitative data in order to facilitate analysis.

Table 1
Summary of Lunar Phase Categories

Category	Qualitative	Quantitative	Mixed	Total
Instrument Development	0	15	3	18
Alternative Conceptions	4	11	3	18
Teaching Approaches	4	1	8	13

This chapter presents a review of the research related to the use of computer simulations in space science. The study of computer simulations to aid in the teaching of chemistry, physics, and biology are specifically not included because these simulations are mimicking either the physical world (such as chemistry and physics) or small-scale natural world (such as biology). Studies related to the simulation of phenomena in these subjects are excluded because they are unlike the large scale observations necessary to understand the phases of the moon.

For the space science studies, there were two groups that produced several related investigations. The space science studies were generally based on interview data that was categorized and transformed into quantitative data, so most were designated as “mixed” method. The mixed methods additionally often include more information about the subjects as do the purely quantitative studies. The qualitative studies generally used case studies in which no attempts were made to quantify findings. The quantitative studies generally used paper and pencil instruments to measure conceptual understanding for specific topics.

There were equal numbers of research papers related to alternative conceptions documentation (18) and instrument development (18) which were both more prevalent than teaching approaches (13). This shows a deficiency in investigations into teaching approaches. The research approaches were mostly quantitative (27) or mixed (14). Studies that were strictly qualitative were not very prevalent (8). Some of these studies specifically studied computer simulations related to space science, including lunar phases (15). These computer simulation studies used mostly a mixed method approach (11), compared to qualitative (3) or strictly quantitative (1) as shown in Table 2.

Table 2
Space Science Computer Simulation Studies by Type

Qualitative	Quantitative	Mixed	Total
3	1	11	15

The following chapter is a review of this literature and is subdivided into the following sections: instrument Development, alternative conceptions of Lunar Phases, effectiveness of instructional approaches, computer simulation examples from space science education, literature review summary, followed by recommendations for research.

Instrument Development

The development of validated assessment instruments is critical to learning about how best to teach lunar phases to students. There have been several different instruments developed in recent years. There have been eighteen studies that have led to the development of either paper and pencil forced-choice instruments, or to open questioning approaches to that are able to better determine the students' alternative conceptions.

There is evidence that shows the more open questioning methods have advantages in determining conceptual understanding of students related to lunar phases. Time will tell if this will be a true paradigm shift in the assessment of lunar phase education, or if it will be simply an addition to established quantitative assessment tools that were already in use.

Hufnagel et al. (2000) and Hufnagel (2002) describe the development of the Astronomy Diagnostic Test 2.0 (ADT 2.0). Hufnagel et al. (2000) provides much more details into the development of the ADT 2.0 but Hufnagel (2002) provides more insight into the evolution of individual questions from version 1.0 to version 2.0 and Hufnagel (2002) was written when more data was available. The cornerstone of the development of this instrument was the assumption that "…student learning during an introductory astronomy course is closely linked to prior knowledge and beliefs about the universe…" (Hufnagel, 2002, p. 47). Three of the 21 astronomy questions were related to phases of the moon. One of the three questions dealt with eclipses and was vocabulary-driven. The other two questions dealing the phases of the moon were related to observations taken from the earth and were conceptual in nature. Since only 3 out of 21 questions were related to lunar phases, the ADT 2.0 can be considered a weak instrument when measuring conceptual understanding of lunar phase concepts specifically.

Zeilik and Morris (2003) reported findings from their study that looked at conceptual gains for college astronomy students that are enrolled in a class specifically for science majors. Previous studies by Zeilik including (Zeilik et al., 1999), (Zeilik, et. al. 1998), and (Zeilik & Bisard, 2000) studied "Astro 101" classes which were specifically designed for non-science majors. All four of these studies are similar in scope and structure, so the most recent work by Zeilik will be addressed. This study was not only looking for alternative conceptions in astronomy, but alternative conceptions that may have held over from these previous physics courses. Significant gains were measured for one lunar phase concept measure, but no significant gain for two other lunar phase measures.

Deming (2002) provided more details about the use of the ADT 2.0 that were lacking in previous studies that used the ADT 2.0, such as Hufnagel et al. (2000) and Hufnagel (2002). The ADT 2.0 National Project was conducted in order to address reliability and validity of the instrument. There were 5346 students that participated in the pre-course ADT 2.0 and 3842 students that participated in the post-course ADT 2.0. The ADT 2.0 national sample resulted in a pre-course mean of 32.4% and a post-course mean of 47.3%. There should have been some explanation as to why these numbers where different. Another approach would have been to have used matched pairs if the goal was to look for conceptual gains in the courses. With all the unknowns about why the two sample sizes are different, the pre-course and the post-course results can not be compared in any meaningful way.

In a similar vein to the ADT 2.0, a Lunar Phases Concept Inventory (LPCI) has been developed (Lindell, 2003; Lindell & Sommer, 2003; Lindell & Olsen, 2002; Lindell, 2001). Lindell (2001) and Lindell and Olsen (2002) documents the establishment of reliability and validity of the LPCI. Lindell (2003) and Lindell and Sommer (2003) document the use of the LPCI to assess college students' conceptual understandings of the phases of the moon. Lindell (2001) documents early development of the LPCI through interviews as well as quantitative measures. An initial 14-question quantitative diagnostic tool to inventory alternative conceptions regarding the lunar phases was developed. Lindell and Olsen (2002) documents the establishment of reliability and validity of the LPCI. At the heart of the LPCI is the recognition that ideally the determination of a student's pre-existing mental models regarding the phases of the moon should be determined by interview. Since this is not practical for large enrollment classes, a multiple choice instrument was the developed which made use of model analysis theory (MAT) to determine an individual's mental model as well as patterns in mental models. The MAT mental models can be represented as a series of "concept dimensions". Lindell and Olsen (2002) discuss the MAT process, which is used to establish reliability. There were eight concept domains within the same single construct of lunar phases. The Cronbach's alpha was 0.75 for the posttest and 0.55 for the pretest. The posttest value was determined to be an acceptable reliability, but the 0.55 for the pretest was considered to be a low reliability. This low value of the pretest was attributed to possible guessing by the students, which makes the LPCI a potentially problematic instrument.

A forced-choice instrument is limited in addressing alternative conceptions by the number of distractors used. If a student does not find their alternative conception in the distractors, they are left to simply guess. Vosniadou, Skopeliti, and Ikospentaki (2004) conducted research to determine if there were differences between instruments that used a forced-choice approach compared to an open response approach. Vosniadou et al. (2004) state their research goal is to compare forced-choice and open methods of questioning using 3-D models. The participants consisted of 3rd grade school children from the country of Greece. The authors found that the forced-choice interview approach showed higher scientifically correct answers compared to the open questioning interview approach, but the forced-choice also exhibited a lower internal consistency compared to the open response. The authors attribute these differences to the forced-choice approach inhibiting the children from making internal models of the earth-moon-sun system.

There is a problem with the forced-choice results being internally consistent therefore a more open format interview should be used. To this end, Trundle, Atwood, and Christopher (2002) developed a semi-structured interview that used both the construction of 2-D as well as 3-D models that can accommodate many different alternative conceptions that students may have. The rationale for using a 3-D model is that 2-D representations can reinforce misconceptions. Unlike a forced-choice, such as a multi-choice instrument that only has three distractors, this qualitative approach allows the list of alternative conceptions to grow out of the research. The interview consisted of a set of scripted questions with follow up questions were asked as needed and depended the responses of the participants. At all stages the researcher asked probing questions to better understand their conceptual understanding. The video tapes were later analyzed and coded as to the participant's level of scientific understanding of lunar phase concepts. The researchers used inductive data analysis to develop a system of patterns and themes so that each interview could be categorized by their level of scientific understanding. This methodology provides an opportunity to obtain rich information regarding the students' knowledge of the causes of the lunar phases.

Bell and Trundle (2008) describe a similar methodology of using scripted interviews, with the addition of a lunar phases card sorting that is completed by the participant at the end of the interview. One disadvantage to the semi-structured interview is that it is more time consuming to administer than a forced-choice instrument. However, a major advantage to this methodology is that participant's responses are not limited to just a few responses of a forced-choice instrument. This results in a richer and more detailed understanding of the participants' conceptions of lunar phases.

Trundle et al. (2010) extend the development of the semi-structured interview protocol by developing a scoring rubric for use with the interview protocol so that a numeric point value (0-10) can be assigned to an interview. A fully scientific understanding results in being assigned 10 points, whereas an interview where the participant has no conceptual understanding of lunar phases results in 0 points being assigned. This rubric is important because the interview data can be used in statistical studies.

In summary, various instruments exist for assessing lunar phase concepts. Each type of instrument has limitations. A forced-choice instrument can not address the many alternative conceptions that students may have and thus suffer from low internal reliability because students resort to guessing. The forced-choice instruments have the advantage that they can be administered to a large number of students and require no specific preparation to administer, unlike semi-structured interviews. An interview can lead to rich information about a student's conceptual understanding, but is difficult to make use of the instrument on a scale larger than a single classroom. Since an interview protocol that asks probing questions allows for follow up by the research, this approach is more likely to give an authentic measure of the student's conceptual understanding. The use of 3-D models with the semi-structured interview avoids any alternative conceptions fostered by a 2-D illustration of lunar phases on a page. However, the forced-choice instruments might be the only practical instrument in situations where not much time is available for assessment.

Alternative conceptions of Lunar Phases

Since the moon is an object that is identifiable by casual sky observers, it can be expected that there exist many alternative conceptions about the scientifically correct explanation of the lunar phases. Six studies investigated alternative conceptions of k-12 students, and ten studies that researched college-age students. What might not be expected, is that many of the misconceptions, such as phases being caused by the Earth's shadow, is held by students at all grade levels. What follows is a review of the studies that specifically identified lunar phase misconceptions.

Alternative conceptions Research at K-12 Grade Level

Jones, Lynch, and Reesink (1987) studied alternative conceptions of lunar phase concepts by studying concepts directed related to lunar phases. The analysis of the concepts students had related to shapes showed that some students were not able to distinguish between apparent and real shapes of especially the moon. Although Jones et al. (1987) was not specifically making categories for the phases of the moon concepts, the phases were discussed by the analysis of the apparent versus real shape of the moon.

Basic research into children's understanding of astronomical events continued in the 1980s in works such as Baxter (1989). It is common that studies do not cite other very similar works. An example of this is Baxter (1989) makes no reference to Jones et al. (1987) even though they were both published in the same journal. Even so, the two studies are very similar in that they document alternative conceptions held by children by addressing very basic ideas such as the sun-earth-moon relationship. Both works use children's previous experiences viewing the sky to assess their preconceptions. The alternative conceptions found that are related to phases of the moon include clouds covering the moon plus three variations of a shadow being cast onto the moon. This study showed that ideas that clouds covering the moon are not held by students older than 12 years. The most prevalent conception is that the shadow of the earth is being cast onto the moon. And to a lesser extent, the correct conception of angles that the moon is being viewed is held. It is noteworthy that the mental model of the earth's shadow being cast onto the moon is not too surprising, since most treatments of the topic do not include any discussion of moon orbiting at a 5 degree angle with respect to the plane of the earth and sun.

Vosniadou and Brewer (1994) continued the theme of Baxter (1989) in that general mental models of young student were investigated. Even though Vosniadou and Brewer (1994) did not mention lunar phases in the title of their work, a significant portion of this study involved concepts of movement of the moon in the sky. This study was well-grounded in the literature and built upon conceptual-change models for mental models related to the teaching of science developed by Vosniadou (1994). An interesting result of this study as a whole is that students often begin with naïve understandings (such as clouds blocking the sun) for the day/night cycle in early grades such as first grade, to be followed with a synthesis of ideas (such as the moon and earth revolve around the sun every 24 hours) in subsequent grade levels. This synthetic mental model is replaced in some students with the scientific view by the end of their elementary school years. This is attributed to the acquisition of knowledge by the children using the conceptual change model as a mechanism for the change in mental models. This change in mental models can be a rough framework for changing mental models of lunar phase concepts.

Stahly, Krockover, and Shepardson (1999) also continued the theme of Baxter (1989) and studied the alternative conceptions that third grade school students specifically have regarding phases of the moon in greater detail. Similar to Vosniadou and Brewer (1994), Stahly et al. (1999) studied how a student's mental model of lunar phase topics can change over time. In general, the 3rd grade students had difficulty with matching the appropriate lunar phase with various earth-moon-sun orientations. The students also can maintain some of their original concepts even after instruction. The students specifically had difficulty with incorporating the concept of eclipses and the earth's rotation into a mental model of the lunar phases.

Hermann and Lewis (2003) attempt to build on the work of Stahly et al. (1999) but their study had many weaknesses. The instrument itself was severely flawed because at least one of the "correct" answers for the multiple-choice instrument was in fact in error. The findings showed that the alternative conception that the earth's shadow causes the phases of the moon is prevalent before instruction and that over one third of the high school students studied hold this alternative conception after instruction. However, there many problems with Hermann and Lewis (2003) and their results should be interpreted with care.

Trundle and Troland (2005) studied lunar phase alternative conceptions from a new perspective. Their research results fell short of finding a direct link between alternative conceptions illustrated in children's literature and the alternative conceptions that children hold regarding the phases of the moon. However, Trundle and Troland

(2005), showed that there are indeed commonalities between alternative conceptions in children's literature and actual alternative conceptions that young students hold.

Trundle et al. (2010) showed that there can be substantial improvement in lunar phase conceptual understanding of middle school students when conceptual change is used to address alternative conceptions. There were two previously unreported alternative conceptions that included: that the Sun's shadow in conjunction with cloud cover causes the moon phases and that the Earth's tilt with respect to the moon causes the moon phases. In each of these cases, only 1 out of 20 participants responded in this way, so neither were significant individually, but contributed to the overall number of alternative conceptions, especially for the pre-treatment results.

The literature on alternative conceptions indicates that the most prevalent alternative conception found for the K-12 level was that the moon was passing into the Earth's shadow to create the lunar phases. Another common alternative conception is that the Earth's rotation on its axis causes the phases of the moon. These alternative conceptions appear to stem from a general lack of understanding about the how the Moon orbits the Earth and in turn how the Earth rotates and also how the Earth orbits the Sun. Other common alternative conceptions include the moon's position relative to geographic areas on the Earth as well as clouds obscuring the view of the moon, thus causing lunar phases. These studies can be useful future researchers, but should they be carefully considered in some cases because some of these studies appeared to introduce alternative conceptions due to lack of rigor in instrument development. In general, understanding the nature of alternative conceptions is important because the conceptual change approach is generally effective for teaching lunar phases.

Alternative conceptions Research at College Level

Since lunar phase concepts are incorporated into many national science education standards for primary and secondary education, there might be a difference between alternative conceptions held in the K-12 grades compared to college level. College students might hold fewer alternative conceptions related to lunar phases than at earlier grade levels. There are no longitudinal studies comparing earlier grades to college level, but useful information might be gleaned from comparing various studies at different grade levels.

One of the earliest quantitative works studying post-secondary students was in 1963 by James Kuethe at Johns Hopkins (Kuethe, 1963). Kuethe was interested in determining “sophisticated errors” in teaching science. These “sophisticated errors” referred to by Kuethe are known in more modern literature as “alternative conceptions”. The only question related to lunar phases resulted in 70% of the students thinking that lunar phases were caused by Earth’s shadow. One of the major weaknesses of this study is that it only looked at male students.

Parker and Heywoood (1998) conducted a study to see how astronomical subject knowledge, including phases of the moon, of pre-service, as well as in-service, teachers can best be put into practice in the classroom in the UK. This study seems to have been a useful tool to improve teaching practice of these particular teachers, but it was weak in that it did not use a formalized post-assessment. It was not apparent if there was specific conceptual change with the instrument that was used.

Suzuki (2003) gave an interesting insight into pre-service teacher preparation in Japan with a pair of case studies. The senior project was to explore approaches to teach the inclining (reclining) crescent moon to the next group of sophomores. It was discovered that pre-service teachers have a difficult time explaining this phenomenon. The major finding was that peer collaborations (for the sophomores) result in vigorous discussions and student involvement is supported by the data provided.

The documentation of alternative conceptions in astronomy is a reoccurring theme within the astronomy education literature. Trumper (2000) devoted an entire study to documenting alternative conceptions on a variety of astronomical topics, including phases of the moon. The research goal was to find alternative conceptions in astronomy that students have specifically at the college level. The instrument that was used was an alternative conceptions measure test that contained 19 questions taken from three previous studies and combined. The questions were multiple-choice and graded for correctness. The question dealing with the cause of the phases of the moon was answered correctly by 51.3 % of the students, which was better than previous studies. One alternative conception discovered was that 71.1% of the students responded that the moon must be "full" in order to have a solar eclipse. Trumper (2000) had many weaknesses, namely the development of the instrument was not fully described since few details were given.

Schoon and Boone (1998) studied the connections that scientific alternative conceptions held by pre-service teachers had with the participant's self-efficacy. It was quite similar to two other studies by Schoon. Schoon (1992) explored alternative conceptions about space science in a mixed sample of 1213 elementary through college-aged students. This research led to more research by Schoon, namely (Schoon, 1995) where the study of alternative conceptions in science, particularly space science, was conducted with 122 pre-service teachers. It was found that pre-service teachers, prior to instruction, have many of the same alternative conceptions as their future students. In both of these early studies, the most common alternative conception relating to the phases of the moon was that the phases are caused by the earth's shadow being cast onto the moon.

There have been attempts to assess expected alternative conceptions over a variety of ages, including college age, within the past two decades (Bisard, Aron, Francek, & Nelson, 1994). These researchers wanted to first investigate and assess suspected alternative conceptions and secondly they wanted to track how these conceptions can change with age. This study found that percentage of correct responses increased as education level increased as the correct response rate increased from middle school to advanced college level. It was also found that pre-service elementary teachers seem to have about as many alternative conceptions as the middle school students. When all responses for the question on lunar phases was tallied, it was found that nearly 60 percent of the students thought that the Earth was involved somehow in producing the lunar phases and only 40 responded that lunar phases were produced by reflected sunlight.

Lindell and Sommer (2003) studied college students to investigate the probabilities of certain alternative conceptions regarding the phases of the moon. The Lunar Phase Concept Inventory (LPCI) was administered to 766 participants prior to instruction in a college-level astronomy course. Analysis showed that the participants were using the correct model for the period of lunar phases, but the incorrect models were "Earth/moon/sun and observed phase relationship; the phase, time of observation and location in the sky relationship and the cause of the lunar phases and no dominant model for the dimension of the moon motion" (Lindell & Sommer, 2003). This research showed that not all lunar phase concept dimensions are equally held prior to instruction.

Lindell (2003) also studied college students enrolled in astronomy courses, but this time the LPCI was used to measure differences in alternative conceptions before and after instruction. The instrument used was the LPCI, which is described above. For pre-instruction, the participants held majority alternative conceptions for five out of eight lunar phase concept dimensions. These dimensions for pre-instruction include: 3.Direction of moon's orbit, 5. Phases caused by sun-earth-moon positions, 6.The phase relating to location in the sky and time, 7.Cause of lunar phases, 8.Effect of lunar phases with geographic location on earth. Six weeks after instruction the LPCI was administered again. This time, only two out of eight dimensions were held by a majority of the participants. These dimensions for post-instruction include: 6.The phase relating to location in the sky and time and 7.Cause of lunar phases. These two dimensions were found to be the most deeply rooted alternative conceptions in the pre-test, so it is not surprising that they have not been moved from the majority held status.

Trundle et al. (2002) studied 63 preservice elementary teachers to measure their lunar phase conceptions before and after inquiry-based instruction. The most prevalent alternative conceptions concerning reasons for lunar phases included: Earth's shadow on moon, Earth's rotation on its axis, and the Moon's position related to locations on Earth. Almost all the students studied held alternative conceptual understanding of the causes of lunar phases before instruction. Trundle et al. (2002) also found that without instruction, pre-service teachers generally hold non-scientific views of phases of the moon. The most common alternative conception was that the lunar phases were caused by shadows cast by the earth (13 out of 42, or 31%). Other alternative conceptions were also discovered including: earth's rotation (5 out of 42, or 12%), moon's relative to position to different geographical locations on earth (2 out of 42, or 5%), as well as atmospheric clouds (1 out of 42, or 2%). The 2-D drawings made during instruction seemed to reinforce the alternative conception that the moon, earth and sun are within the same plane, which was thought to be the root cause for the alternative conception that phases are caused by the earth's shadow. Another finding was that several participants that were classified as having a scientific understanding, upon further questioning indicated that the earth's rotation on its axis contributes to the phases of the moon. It was speculated that the problem stems from the psychomotor model activity that was used for the 3-D modeling. In the 3-D models, the participant's head was used to represent the earth, a light source was the sun, and a ball was the moon. As the participant rotated, the lunar phases were simulated on the ball. The use of the 3-D models appeared to have advantages over the 2-D models including richer responses from participants.

Trundle and Bell (2010) showed that by far the most prevalent alternative conception for the cause of lunar phases was the Earth's shadow on the moon. Moreover, Trundle and Bell indicate that this result is consistent with 13 other studies that included over 3000 participants.

In general, the research indicates that few college students hold a scientific conception for the lunar phases prior to instructions. These results were very similar to alternative conception research that studied K-12 students. This is echoed in Trundle et al. (2010). The most common alternative conception for both age groups was that the shadow of the Earth is being cast onto the Moon to cause the lunar phases, also can also be referred to the eclipse explanation.

Effectiveness of Instructional Approaches

The preceding studies that investigated alternative conceptions that student hold regarding lunar phases set a foundation to develop new teaching strategies for lunar phase concepts. By specifically addressing misconceptions, the student's conceptions can be changed using the conceptual change model of science instruction. There were ten studies that mostly investigated the conceptual change approach to lunar phase instruction. Trundle, Atwood, and Christopher (2007a), drawing on their previous work Trundle et al. (2002, 2006) as well as the work of others such as Vosniadou et al. (2004) develop perspective of conceptual change that moves beyond descriptive studies of alternative conceptions to using this knowledge of alternative conceptions to specifically address them during instruction. The researchers can now start making use of the conceptual change model along with an inquiry-based mode of instruction.

The assessment of teaching approaches for lunar phase concepts is less studied than alternative conceptions discovery or instrument development. A unique approach was reported by Sadler (2000), in which observational journals were a major component of the course. This study was mostly a non-empirical work that describes the approach taken at a specific university for a specific course. Although there was not any comparison of teaching approaches in this work, it was informative to see how journal keeping can be used in a college astronomy course.

Abell, George, and Martini (2002) used a pre-service science teaching methods course to not only assess the student learning among college students but to affect their teaching by a series of reflections made during the learning process. This study was similar in many ways to Parker and Heywood (1998) in that teacher reflections of their own science learning can positively affect their own science teaching.

Trundle, Atwood, and Christopher (2002, 2006) provide an opportunity to investigate a class of pre-service teachers that only used inquiry-based lunar phase instruction with real sky observations. Trundle, Atwood, and Christopher (2007a) is similar, but investigates students in the fourth grade. Trundle, Atwood, and Christopher (2007b) is also similar, but is a longitudinal study that studies conceptual understanding of pre-service teachers several months after instruction. All of these studies also used the manipulation of 3-D objects during instruction in addition to students keeping journals. They all showed that inquiry-based instruction can result in substantial conceptual gains in the scientific understanding of lunar phases. Bell and Trundle (2007, 2008) and Bell et al. (2007) also employed inquiry-based instruction that used 3-D models and journal keeping, but in contrast with other studies, they made use of *Starry Night*© planetarium software to simulate real sky observations. These studies also showed a dramatic increase in scientific understanding of lunar phases after instruction. In some cases, the use of *Starry Night*© showed a significant improvement over real-sky observations. A more detailed discussion of the use of planetarium software in the classroom will be presented in a subsequent section.

Trundle et al. (2010) showed that for middle school students, the conceptual understanding of lunar phases can be greatly increased with the use of guided inquiry instruction when gathering moon phase data from the real sky.

Trundle and Bell (2010) used three distinct treatments: *Starry Night©* observations only, *Starry Night©* and Natural Observations, as well as Natural Observations only. The three treatments resulted in substantial and significant conceptual gains when comparing results after instruction to results before instruction. However, the same study showed no statistical differences between treatment groups for the participants' ability to draw scientific moon shapes or their conceptions of the causes of lunar phases. Additionally, the study showed that the use of planetarium software, namely *Starry Night©* in conjunction with an inquiry-based, conceptual change approach, resulted in statistically greater gains in sequencing lunar phases when compared to other treatments that observed that real night sky.

In general, few studies compare various teaching approaches, although the conceptual change approach to teaching lunar phases has shown significant improvements after instruction. Most of the studies lack detail of their instrument and simply looked for "conceptual growth" of students (Barnett & Moran, 2002) or the thoughtfulness of student reflections (Abell et al., 2002). Only research by Trundle and her colleagues provide details of their studies that used a standardized instrument to measure gains conceptual understanding of lunar phases so that different teaching approaches could be compared. Overall results suggest that the use of planetarium software to simulate the night sky, such as *Starry Night©* in conjunction with an inquiry-based, conceptual change approach holds the most promise as being the most effective approach to teach lunar phases. Moreover, *Starry Night©* software can promote a constructivist approach to science education because the participants were constructing their own knowledge by using conceptual change.

Computer Simulation Examples from Space Science Education

In the development of computer simulations for observational astronomy activities, two groups have emerged. Those that develop dedicated educational software in the same way many other researchers have in the fields of Earth Science, Biology, Chemistry, and Physics. The other group that has emerged uses commercially available planetarium software to simulate the motions in the night sky. Astronomy educators benefit from a robust market for planetarium software for the amateur astronomy enthusiast. This makes very powerful planetarium software available at very reasonable costs, especially when compared to the cost of developing dedicated software. However, there are some efforts to develop dedicated computer software to promote inquiry-based, constructivist approaches to teaching of space science concepts.

An example of development of dedicated software for night sky simulations is the "pictorial computer simulation" as described in Kangassalo (1994). The main finding of this study was that the more developed a child's conceptual model, the more they investigated the computer simulation with a specific aim.

Another example of the development of dedicated software for night sky simulation is the Virtual Solar System (VSS) project as described in (Barab, Hay, Barnett, & Keating, 2000; Barab, Hay, Squire, Barnett, Schmidt, Karrigan, Yamagata-Lynch, & Johnson, 2000; Keating, Barnett, Barab, & Hay, 2002; Gazit, Yair, & Chen, 2005). Barab, Hay, Barnett et al. (2000) establishes the rationale for the development of the VSS. At the heart of their approach is the constructivist approach in which students construct their own meaning of phenomena. In their "nested relationship" model, Barab, Hay, Barnett et al. (2000) surrounds the constructivist approach with project-based assignments. These project-based assignments are then surrounded by the participatory engagement of students. The students used a virtual reality modeling language (VRML) to develop an immersive virtual solar system (VSS). The findings presented in Barab, Hay, Barnett et al. (2000) suggested that students can develop rich understandings of solar system phenomena. Barab, Hay, Squire, et al. (2000) is related to Barab, Hay, Barnett, et al. (2000) and in fact uses some of the data that is described in Barab, Hay, Barnett, et al. (2000). Barab, Hay, Squire, et al. (2000) is different in that more analysis of made of the data with less reliance on the body of literature. Keating, et al. (2002) was able to build on the data set originally described by Barab, Hay, Barnett, et al. (2000). Keating, et al. (2002) report findings from pre and post interviews with the students to quantify gains in understanding astronomical concepts, and the second goal was what type of conceptual understanding does the 3-D modeling facilitate. Overall the researchers found that the VSS course demonstrated statistically significant gains in conceptual understanding for reasons for the seasons, eclipses, and lunar phases. Through the interviews, the researchers were also able to determine that "3-D computer modeling software not only allows students to place themselves in the role of the Earth, Sun, a third person observer, or the Moon, but also provided them with a powerful too to test and revise their models."(Keating et al., 2002, p. 272).

Hanson, Barnett, MaKinster, and Keating (2004) continue the analysis of the data first described by Barab, Hay, Barnett et al. (2000). The VSS course used the same three projects as described in Barab, Hay, Barnett et al. (2000). The goal of this study was to look for differences in types of learning by quantitatively comparing the VSS and a traditional lecture-based astronomy course. Their findings show that show that VSS shows higher scores over the traditional lecture approach for concepts related to spatial understanding of solar system object. The traditional lecture approach to teaching was found to best facilitate fact-based knowledge.

Barnett, Yamagata-Lynch, Keating, Barab, and Hay (2005) continue the investigation in using the VSS. This study was very similar to other studies in that a relatively small group of college students was studied in their use of the *Virtual Reality Markup Language* (VRML) in the VSS course. Barnett et al. (2005) find that computer modeling supported students learning concepts that require a change in the frame of reference or the perspective of the observer.

Gazit, Yair, and Chen (2005) investigate the use of the VSS approach with 10^{th} grade students. Unlike previous studies that investigated VSS courses, this study looked at students in a laboratory environment and studied their activities in great detail. Also unlike other VSS studies, the students worked alone. Each student was studied over a series of four 1.5 to 2.0 hour sessions. In general, there was found to be more interactions with the 3-D model for the more complex questions, such as the lunar phase question. Another finding was that students make use of the drawing tool on the computer screen. The interviews also found that alternative conceptions can be developed due to the learners' difficulty dealing with different frames of reference as well as difficulty with visualizing the 3-D nature of the simulation. The authors suggest careful scaffolding of the use of the software and guided reflection to minimize the growth of alternative conceptions.

Bell and Trundle (2006, 2007, 2008) report a continuing study that makes use of commercially available software for simulation of the night sky. This study used *Starry Night*© which is a brand of planetarium software. This type of software can mimic the view of the night sky and gives the user control over the night sky much like extensive planetarium systems, but with the use of inexpensive personal computers. The initial concept of using planetarium software for teaching lunar phase concepts was reported in Trundle and Bell (2003).

For the intervention, the students used *Starry Night*© to make daily "observations" of lunar phases during a nine week period. The students then used these observations, along with hands-on activities and class discussions based on *Physics by Inquiry* by Lillian McDermott (1996), to confront their alternative conceptions about lunar phases and develop a more scientifically correct understanding. Pre and post instruction interviews were conducted and these interviews were scored according to the students' scientific understanding of lunar phases. The assessment consisted of a semi-structured interview protocol in which students completed six different tasks as described in Trundle et al. (2002).

Bell and Trundle (2008) find that none of the participants held a scientific view of lunar phases before instruction, but 82% of the students held a scientific view of lunar phases after instruction with *Starry Night*© only. Bell and Trundle (2007) in addition to the findings above, also show that when students use *Starry Night*© in conjunction with natural observations, the scientific understanding of lunar phases changes from 3.3% before instruction, to 67.2% students having a scientific view of lunar phases after instruction. Additionally, using natural observations only, the percentage of students holding a scientific view of lunar phases started with 2.2% for pre-instruction, and changed to 71.7% for post-instruction. So, using *Starry Night©* alone shows a greater conceptual gain than either approach. Bell and Trundle (2008) provide possible causes for improvements in conceptual understanding by using only the *Starry Night©* approach. These reasons include: reduction of complexity of the observations, more accurate measurements are possible, as well as the lack of interference from weather. Indeed, *Starry Night©* can likely help to reduce many sources of interference that occur when the natural world is observed. This was also true with Winn et al. (2006) that showed that a computer simulation for oceanographic concepts can provide experiences that are impractical when making observations/measurements in the real natural world. Bell and Trundle (2008) report that several students reported that they intended to purchase *Starry Night©* themselves to use in their own classrooms. This gives the *Starry Night©* approach an advantage over the VSS approach, which is not widely available commercially and requires much more extensive preparation to use.

Bell, Binns, and Smetana (2007) studied forty-nine ninth graders. Of those studied, 27 made natural lunar phases observations and 22 used *Starry Night©* software to simulate observations. Both treatments were effective in achieving conceptual change of lunar phase concepts, with the *Starry Night©* treatment being the most effective. The *Starry Night©* group had a 71% gain in those that had a scientific understand of lunar phases, while the natural observations group had a 44% gain. This study complemented Trundle et al. (2006) as well as Bell and Trundle (2008) quite well, because it extended their findings to a younger age group and that it also found that using *Starry Night©*

alone can result substantial gains in conceptual understanding of lunar phase concepts. Bell et al. (2007) also discovered that the LPCI overestimates conceptual understanding of lunar phases compared to the interview protocol because of student guessing with the forced-choice instrument. The LPCI provides perhaps a problematic measurement because it did not discover that the *Starry Night*© treatment was the most effective. Further investigation of the LPCI is warranted. However, since the LPCI so easy to administer it should be still reconsidered and compared to the more-lengthy interview protocol results.

Trundle and Bell (2010) showed that the use of *Starry Night*© resulted in statistically greater gains in sequencing lunar phases when compared to other treatments that observed that real night sky. Although the same study showed no statistical differences between treatment groups for the participants ability to draw scientific moon shapes or their conceptions of the causes of lunar phases.

Many of the studies examined have lacked consistent instruments to evaluate new teaching approaches. Without a consistent evaluation instrument, it is problematic to draw conclusions as to the effectiveness of various interventions. Trundle et al. (2002) used a well-designed interview-based instrument to assess students' conceptual understanding regarding lunar phases before and after instruction that was later adopted by Bell and Trundle (2008) and Trundle and Bell (2010). It was found that students' conceptual understanding of lunar phases can significantly increase when using *Starry Night*© software in conjunction with an inquiry-based real sky observation assignment. Similar findings were also reported in Hobson et al. (2010).

Smetana and Bell (2012) in a critical review of literature relating to the use of computer simulations in science education can facilitate conceptual change. Although this review studied a wide variety of papers from many disciplines, it was found that overall that simulations can be as effective, and sometimes, even more effective than traditional instruction. Moreover, it was found that computer simulations are most effective when they are used to supplement, not replace, other instructional approaches. This finding was echoed by Rutten et al. (2012) who, in a similar critical review, found that computer simulations can enhance traditional instruction, especially when used to simulate laboratory exercises.

There were no studies specifically relating to lunar phases that investigated the issue of instructor-controlled simulation versus student-controlled simulations. However, Smetana (2008) showed that there was no significant difference for a whole-class setting versus students working in small groups at computers for topics in chemistry among high school students. However, for earth science achievement, the teacher directed computer simulation was more effective than student-controlled computer simulations among high school students (Chang, 2003).

The use of computer-based simulation software in order to reproduce observations and simulations of space science events has been shown to be effective. The virtual solar system project, in which virtual reality modeling software was used to model the solar system, was found to be an effective approach to teaching solar system concepts. However, VSS can be difficult to implement because of the high cost. The commercially available and relatively low-cost software *Starry Night*© was found to be quite effective in teaching space science topics, namely the phases of the moon. Additionally, there are initial indications that the use of *Starry Night*© alone can result in greater gains in scientific conceptual understanding of lunar phases compared to real sky observations or even a combination of both observation approaches. The issue of instructor-controlled versus student-controlled simulations has not been investigated and warrants further study.

Literature Review Summary

Lunar Phases

The study of teaching phases of the moon to students has resulted in the discovery of three distinct groups of researchers as well as three distinct methodologies used in the collection of data. In general, there has not been much collaboration between those that study each group. The same lack of collaboration can be seen between practitioners of the quantitative and qualitative research methodologies. Three distinct research communities, that study how students learn lunar phase concepts, have emerged from the body of research. One group is the "astronomers" that are primarily studying college students that are taking astronomy to fulfill their science requirements. Another group of researchers are those researchers from education that are studying both pre-service teachers, as well as K-12 students. A third distinct group of researchers that has emerged are researchers from cognitive psychology that primarily study how elementary-aged students learn lunar phases. A common finding among all groups, regardless to methodology, was the pre-instruction level of conceptual understanding of the phases of the moon was poor. Inquiry-based instruction was used in many of these studies and has shown to be effective.

The research that studied K-12 students primarily focused on grades 3-8, and not the high school grades. This is likely due to the curricula for grades 3-8 specifically address phases of the moon (National Research Council, 1996). All of the research consisted of small-scale studies that looked at no more than a few classes from a single school. The two groups of college students studied in the literature were pre-service teachers and general college students. The pre-service teachers were typically enrolled in a science teaching methods course being offered within a school of education while the general students were taking "Astro-101" courses that were generally offered by either astronomy or physics departments.

The different approaches to studying the teaching of lunar phase concepts were qualitative, quantitative, as well as mixed-methods. For the most part, those researchers studying pre-service teachers within a science teaching methods class used qualitative or mixed method approaches. There are some examples of quantitative approaches used as well. The qualitative approaches included semi-structured interviews, journals, and classroom observations. The mixed-method mode involved converting qualitative data to quantitative data during the analysis process. The rich qualitative data lead to researchers knowing the pre-service teachers much better than those that studied the Astro-101 students with quantitative approaches only. The inclusion of qualitative data makes the process of generalizing to other pre-service teacher groups possible, because the reader can gain special insight into the students that were studied. This is lacking in the quantitative approaches that used a forced-choice instrument. The two main quantitative instruments discussed, the ADT 2.0 and the LPCI, both had demographic questions included. The ADT 2.0 suffers as an instrument for measuring conceptual understanding of lunar phase concepts, because only three out of 21 questions deal specifically with lunar phase concepts, whereas the LPCI is solely measuring lunar phase conceptual understanding. Even with certain problems with each of the protocols and instruments, there are now techniques for the assessment of lunar phases conceptual understandings among students, although care must be taken when interpreting results. The availability of these instruments and interview protocols mark a significant advancement compared to only two decades ago.

When comparing the groups of students, there were some things in common and also some differences. They were similar in that the level of pre-instruction conceptual understanding was very low for phases of the moon, although some groups showed fairly large improvements after instruction. A finding common among the studies was that there are several alternative conceptions that reoccur prior to instruction. The most common alternative conception prior to instruction was that the lunar phases are caused by the earth's shadow. This finding also shows that there is also an alternative conception related to the cause of lunar and solar eclipses. Other common alternative conceptions include: lunar phases caused by earth's rotation, lunar phases being caused by geographic location on the earth, and lunar phases being caused by clouds. A common theme to this research is that knowing what student alternative conceptions are prior to instruction, the teaching practice can be improved.

The majority of these studies did not compare the teaching approach to the lunar phases, however a frequently used approach for teaching phases of the moon could be best described as being an inquiry-based approach. These inquiry methods of teaching could be anything from a direct sky observation (Trundle et. al., 2002, 2006) to small group lecture tutorials (Lindell, 2003). In general, the approach would have learners conducting their own investigation and analyzing their own data to build their own conceptual model. Since there are usable assessment instruments and methodologies in place at this time, perhaps more work should be conducted in teaching approaches and methods.

Computer Simulations

There are empirical studies that investigate the effect of computer simulations on the conceptual understanding of Space science concepts. However, some works such as MacIntyre (1999) and Trundle and Bell (2003) call for the use of planetarium software to be used in the classroom, but lack empirical data.

Studies from other fields can be used to inform the use of computers to teach space science concepts. Monaghan and Clement (1999) showed that computer simulations can facilitate students making mental models for the physics concept of relative motion among high school students.

Of the empirical studies that were critically reviewed nine investigated space sciences simulations. The studies of computer simulations of space science topics show that these simulations are generally effective and show improvements over a traditional large enrollment classes that used a passive lecture format. However, the virtual solar system project is also likely not to be used widely, since the 3-D modeling software and excessive cost seems to be the main deterrent for wide scale use. The studies that investigated the virtual solar system project were generally sound, but they failed to control for the variable of class size. Many of the virtual solar system studies compared learning outcomes of a small class (less than 10) to the outcomes of a large (around 70) lecture class. The studies that investigated the use of *Starry Night©* software did a better job of controlling for variables since groups of participants were roughly the same size for each treatment. Also, control groups were established so that the conceptual gains when using *Starry Night©* could be quantified above and beyond that of the control group. *Starry Night©* also has the added benefit that it is widely available at a relatively low cost.

There is no research related to the issue of computers being used as a demonstration for the whole-class versus student-controlled computer simulations. Again, research from other disciplines can be used to inform the teaching of lunar phases. The results have shown that the results have been mixed and further research is warranted.

The incorporation of simulation software for space science topics was successful, and the approach shows more than just equivalence with natural observations, but show improvements above and beyond natural observations. The most cost-effective software package for space science simulation was *Starry Night©* and it also showed the most improvements in conceptual understanding of space science topics, namely phases of the moon. As for the issue of whole-class versus student-controlled simulations, there is no empirical evidence yet for lunar phase concepts. The literature that studied other science topics showed that the results from whole-class versus student-directed simulations were mixed.

Recommendations

Since inquiry-based instruction holds the most promise, how should this best implemented? Inquiry-based lunar phase instruction could include several months of moon observations using the real sky. However, problems can arise when students make real sky observations. These can include weather, student's schedules, as well as lack of participation of some students when they know that the observations will ultimately be pooled together (Alexander, 2006). Ideally students should be able to gather their own data for science inquiry, but doing this might take more time than is available for a typical course that included lunar phase concepts. However, Chun-Yen (2003) showed that instructor-directed computer simulations were more effective than student-controlled simulations for earth science achievement for high school students. Also, Smetana (2008) showed that there was no significant difference between student-directed compared to instructor-directed simulations for chemistry topics among high school students. Therefore, the issue of whole-class versus small group manipulation of computer software is not yet settled, especially for lunar phase education research.

Computer technology, in the form of night sky simulations, is the key to offering students inquiry-based instruction for lunar phases, while still keeping the time required relatively small so as to include other necessary topics in the course as well. One novel alternative to direct sky observation of the lunar phases is the use of "planetarium software" that simulates the night sky. The use of this software could potentially shorten a lunar observation exercise from a month (or more) to only a few computer sessions. Shortening the observation time will allow more classroom instruction time for other important topics. The observations should also be conducted in a purposeful way that considers how the observations will be used to develop mental models. One way to investigate the reduction of data and to align observations with classroom activities is to take data only during the waxing phases of the moon to see if students can correctly predict the waning phases and thus the entire cycle. The use of planetarium software would also have the added benefit of oversight of the instructor during the activity. This could potentially reduce the frustration and struggle that sometimes can accompany real sky observations (Abell et. al., 2002). The use of planetarium software as a means to facilitate inquiry-based teaching of lunar phase concepts should be explored by the use of randomly assigned experimental designs for classes of pre-service teachers since they are the best-studied group of students in the literature. Such a study could be carried out by gaining access to several sections of the same course during a single semester, or by studying a single section in subsequent semesters. An important consideration would be that the sections should be taught by the same instructor, so that the study is not measuring differences in instructors. Pretest assessments should also be taken of each class so that sections can be measured for equivalency. Care should be taken to make both sections as similar as possible, with the exception of the variable that will be studied, the specific uses of planetarium software.

If we consider the more specific topic of how best to teach lunar phase concepts, it is only natural to build upon the work Bell and Trundle (2006, 2007, 2008) and Bell et. al., (2007). With the success of Bell and Trundle (2007) that showed substantial conceptual gains over and above a traditional inquiry-based approach, it is imperative to encourage wide scale use of this software. One way to make *Starry Night©* more accessible to not only the college classroom, but also the primary and secondary classroom, is the use of *Starry Night©* as a demonstration tool. It would be worthwhile to investigate any differences between a classroom using an inquiry-based *Starry Night©* approach with two students at each computer, and the same inquiry-based *Starry Night©*

approach using a computer projector to provide a demonstration so that the entire classroom can follow along. Ideally the use of a demonstration would be equivalent to each student using their own computer, although it is entirely possible that the conceptual understanding might even be higher with a demonstration. One consideration for the use of *Starry Night©* is the inevitable learning curve that is associated with the new software. If some students are frustrated with this learning curve, their conceptual gains might actually be lower than if they watched a *Starry Night©* demonstration followed with an inquiry-based debrief and model building.

Since *Starry Night©* software shows the most promise to provide inquiry-based space science teaching and that is generally accessible and low in cost its use should be encouraged. Although *Starry Night©* is relatively low in cost compared to other user-developed software, it can still be cost prohibitive and lower-cost alternatives should be considered. For instance, the use of *Starry Night©* as a demonstration would reduce the cost per classroom from $1000, for a computer lab, to $50 for a single use license. The use of *Starry Night©* as a demonstration should be investigated.

It is apparent that the field of computer simulations to promote inquiry-based learning for the earth and space sciences is not yet a mature field. There is still much more research that can be performed. One source of data that has not yet been explored is the investigation of drawings and reflections that students make during instruction. Several studies such as Bell and Trundle (2008) and Bell et al. (2007) investigate the drawings made by students during the interview process but not during instruction. It would be useful to know if drawings made by students during instruction would be useful in predicting their conceptual understanding of lunar phases after instruction. Moreover, it would be of interest to determine the most important factors of instruction as reported by students after instruction.

Research Questions

The following research questions are informed by the recommendations above. They will serve as a guide for the methodology employed to answer them. The research questions are as follows:

1. How do pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases change after inquiry-based, conceptual change instructional intervention based on a computer simulation?
2. Are there differences in achievement between pre-service elementary teachers who collect moon observation data in a whole-class setting versus those who collect this data working in pairs with the computer simulation?
3. How do pre-service elementary teachers perceive the relative value of each component of the instructional intervention as related to their attainment of scientific conceptions of moon phases?
4. How does the forced-choice LPCI instrument compare with the semi-structured Trundle protocol in measuring the conceptual understanding of lunar phases among pre-service elementary teachers?

CHAPTER 3: METHODOLOGY

Methodology

The inclusion of the topic of lunar phases is a critical concept in the preparation of future pre-kindergarten through eighth grade (pK-8) teachers because it is a key concept for space science according to national standards. Research into the teaching of lunar phases is critical to insure quality science instruction into the future. The *Benchmarks for Scientific Literacy (Benchmarks)* state that students need the direct experience with physical models in order to understand the geometry of the sun-moon-earth system (AAAS, 1993). Moreover, the *National Science Education Standards* (NSES) include phases of the moon and earth-moon-sun relationships (NRC, 1996). The NSES stress the importance varying shapes of the lunar phases, the sequence which the occur, the relative positions of the moon, earth and sun, as well as the length of time it takes for a lunar cycle to occur. Revised national standards continue to call for specific instruction regarding lunar phase concepts, for example the *Next Generation Science Standards(NGSS)* (National Research Council, 2013) give specific recommendations for making use of physical models to help students understand the geometry of the sun-earth-moon system that is critical for understanding phases of the moon. Empirical research suggests that an inquiry-based, constructivist teaching approach with students making real world observations of the changing lunar phases and then analyzing their data is an effective way to teach lunar phases. However, making real-world sky observations can be time-consuming, taking up many hours of class time as well as time outside of class (Trundle et al., 2006). Real-world observations of the night sky can also lead to student frustration as well as students not making their own observations to benefit their group discussions (Alexander, 2006). Computer-based planetarium software has shown promise in simulating lunar phase observations and can result in a less time-consuming approach toward instruction (Bell & Trundle, 2008). In past studies the collection of data has ranged from sixteen days to up to nine weeks. No studies have specifically investigated how many observations are required for students to develop a scientific understanding of lunar phases. One approach to investigate an acceptable number of observations is to take data only from new moon to last quarter moon to see if students can correctly extend the patterns they observe to complete the waning phases and thus the entire cycle.

Empirical research also suggests that an inquiry-based, constructivist teaching approach using computer-based planetarium software can be effective in teaching lunar phases Bell and Trundle, (2006, 2007, 2008) and Bell et al. (2007).

There are relatively few studies that investigated students using computer-based planetarium software to make lunar phase observations, such as Bell and Trundle. Moreover, they only study students making their own observations either alone or in small groups of two or three. They did not study classrooms that use planetarium software in a demonstration mode, where the instructor controls the computer and the students observe with the aid of a single computer projector. This could be a beneficial instructional approach since it would be more cost-effective compared to purchasing a site license for a dozen or more computers. This study used a quasi-experimental design to compare the overall conceptual understanding of lunar phases between two groups of pre-service teachers that are exposed to different instructional approaches. In one treatment students made lunar observations in small groups of two using a laptop computer and with the other treatment the instructor operated the planetarium software which is projected by a single projector, while the students make observations. In both cases the instructor provided guided instruction with scaffolding so as to promote conceptual change. For both treatments, the student worked in groups to analyze their data in groups of two. Thus limiting differences of the two treatments to only how data are collected, not how data were analyzed.

In the Trundle studies, students did not keep a journal in which they reflected on their learning. For the current treatment, students gathered their data on data sheets and at the end of each class period wrote a reflection related to their own learning process. A typical student journal might be kept during an entire semester but due to the compressed schedule for the current treatment, only sketches and daily reflections were kept. The student data gathering on data sheets and writing reflections was more structured than perhaps a typical semester-long journal and will included the following elements:

- Drawings made during instruction
- Description of the data gathered
- Possible explanation of the data gathered
- Predictions related to the cause of the lunar phases
- Questions for next class
- Any major surprises from the day

As part of the quasi-experimental design, the daily student reflections from both treatments were analyzed in order to better understand how students develop their understanding of lunar phases as related to the conceptual change approach to science education. Since the two treatments provide an opportunity to compare much different ways for students obtain data for their journal entries, it illuminated the journal-keeping process as a way of constructing an understanding of lunar phases. This present study used the student reflections, as well as post-instruction interviews, to investigate what principle components there are for lunar phase teaching and learning.

Purpose

This study consisted of investigating several aspects related to teaching lunar phase concepts to pre-service teachers that are undergraduates. As part of the research design, the scientific understanding of lunar phases was measured by both a semi-structured interview as well as the LPCI. The LPCI was used for pre and post instruction assessment as well as a screening tool to purposely sample a random sampling of the students for the more extensive interview protocol. The scientific understanding of lunar phases as measured by the interview requires assesses student understanding of the shapes, sequences, as well as causes of the lunar phases as seen from Earth. Student reflections were also analyzed to track the progress of students forming conceptual models of lunar phases.

The present study investigated the use of the conceptual change approach of science education where the learners are able to confront their alternative conception with data they have collected. The students are able to revise and replace their mental models in light of the data they gathered. More specifically, the present study compared students gathering data in pairs to students gathering data as a whole class. The interviews and the student daily reflections were particularly useful to illuminate the process of conceptual change for students learning about lunar phases.

This study specifically addressed deficiencies in the literature by answering the following research questions:

1. How do pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases change after inquiry-based, conceptual change instructional intervention based on a computer simulation?
2. Are there differences in achievement between pre-service elementary teachers who collect moon observation data in a whole-class setting versus those who collect this data working in pairs with the computer simulation?
3. How do pre-service elementary teachers perceive the relative value of each component of the instructional intervention as related to their attainment of scientific conceptions of moon phases?
4. How does the forced-choice LPCI instrument compare with the semi-structured Trundle protocol in measuring the conceptual understanding of lunar phases among pre-service elementary teachers?

Theoretical Framework

The overall goal for this research was to investigate various facets of teaching lunar phase concepts using an inquiry-based, conceptual change approach. To this end, the instructional treatment included students gathering data and using them in constructing a mental model that are used to replace their original mental models based on alternative conceptions using a conceptual change approach. Each student participant confronted their alternative conceptions regarding lunar phases and then built a correct scientific mental model. According to Vosniadou and Brewer (1994) a mental model is built in stages, starting with the initial mental model, then progressing to a synthetic model after some instruction and then finally to a scientific model. Vosniadou and Brewer (1994) go on to develop a perspective of conceptual change that alternative conceptions related to lunar phases can be highly resistant to change. This finding is also echoed by Staver (1998) who calls for conceptual change as well as cooperative learning to address students' tendency to hold on to their alternative conceptions. To this end, participants worked in groups to construct both 2-D and 3-D models of the waxing phases and 7 days into the waning phases which they had observed. The 2-D models were drawn with paper and pencil. The 3-D models were also constructed with the use of Styrofoam® balls (scientific models). Stahly et al. (1999) found that repeated instructional activities facilitated conceptual change for lunar concepts, albeit for third graders. So having the student participants build their mental models in a variety of ways was beneficial and addressed their alternative conceptions.

The models that participants constructed were then tested by making a prediction of the waning phases from last quarter to new moon. The predictive ability of a model is a key item in an inquiry approach to science teaching, followed by experimentation (or observation), formalization and lastly generalization (White, 1993). Therefore the use of data from new to last quarter moon was judged on the predictive ability of the mental models developed. The moon phase of new to third (last) quarter was selected not only due to a relatively short length of observations (21 days), but also because these phases represent exactly three quarters of a lunar cycle. In constructing the models the first three quarters of the lunar cycle was used to predict the shape, sequence, and position of the last quarter. The approach was used to make the lunar observations was based on the assumption that the first three quarters of the cycle could be used to predict that last quarter of the cycle. To this end, treatment A had the participants making observations of the moon during new to last quarter phases by means of a demonstration of *Starry Night*© software conducted by the instructor of the course. Alternatively, treatment B had participants taking their data using *Starry Night*© software working in groups of two at each computer. In measuring each group's ability to predict the waning phases from last quarter to new moon, not only are the two treatments compared, but also the mental models that are developed.

The investigation of student journals kept during the instruction of lunar phases has not commonly practiced among researchers. Many of the studies from science education outside of lunar phase studies present student journal findings that are non-empirical. Ruiz-Primo and Li (2004) showed that assessments of science notebooks are shown to be an accurate indicator of student achievement. Sadler (2000) specifically showed that students keeping observational journals during real sky observations experience a creative process and college students were able to test and improve mental models with the use of their data. In order to better understand differences between the two treatments, measuring any differences with how students develop mental models gave a unique insight.

Vosniadou and Brewer (1994) discuss the concept that mental models are built by the learner in stages through the investigation of different age groups of students and not by the investigation of the learning process of individual students. For the present study, the use of the student reflections gave a unique insight into this process. The reflection data were collected and correlations with other outcomes were investigated using predictive statistics.

Participants and Sampling

The participants for this study were students at a mid-sized eastern university that is primarily an undergraduate institution with a strong program for teacher education. The participants were planning to become pK-8 teachers. The students were enrolled in the course General Science-162 (GSCI-162), whose title is “Science of the Planets”. This course is the second in a five-course sequence that interdisciplinary liberal studies (IDLS) majors take to fulfill their general education science requirements. The courses in the GSCI-16X series are taken only by IDLS majors so the students have similar backgrounds and have similar needs for using the material in the future. All of the students in the classes were majoring in IDLS and are planning on becoming pK-8 teachers, but they have not yet declared their areas of specialization. Approximately one quarter of IDLS majors will go on to declare science as a specialization area. Over 90% of the students in GSCI-162 courses have historically been female and over 95% have been Caucasian with an average age of approximately 19 years. For the present study, only 2 of the 69 participants were male. Additionally, only 2 of the 69 participants represented a minority group (Asian). Therefore, gender and ethnic background differences were not possible to investigate in the present study.

Only courses that are specifically designed for pre-service teachers were studied. The study included four sections of GSCI-162 taught by the same instructor (not the investigator) over two semesters. Two sections each semester received different treatments, as described below. The section sizes are typically 15-20 students. For the present study there were 15 participants for treatment A in the first semester and 19 participants for treatment A for the second semester. For treatment B, there were 16 participants for the first semester and 19 participants for treatment B during the second semester. The total for both semesters for both treatments resulted in 69 participants. All of the students enrolled in the classes agreed to participate in the present study, so there was 100% participation.

Participation in the study was voluntary and the identity of individual students was and will remain confidential throughout the study and afterwards. Student grades for the course were in no way be affected by their participation in the study. Participants were free to leave the study at any time during or after the study. Participants were recruited by the researcher rather than the instructor for the course in order to minimize any perceived coercion. However, as an incentive for students to participate, each participant was given a planisphere which is a tool for night sky observing and has a nominal value. Additionally, a local coffee shop gift card of nominal value was given to each participant randomly selected for the interview protocol.

Treatments

An inquiry-based approach based on *Physics by Inquiry* by McDermott (1996) was used as the basis for both instructional treatments. A major change compared to *Physics by Inquiry* was to use planetarium software, namely *Starry Night*© to simulate the night sky for the observations of the moon as described in Bell and Trundle (2008). The participants gathered lunar data for 21 days (three quarters of a lunar cycle) and then made predictions regarding the last 7-8 days (last quarter of the cycle). Also, the treatments did not specifically cover eclipses as well as time-of-day visibility of specific lunar phases. This instructional approach is similar to Trundle et al. (2007a,b) in that students worked in groups to compare observations, and then worked on physical and mental models of the earth-moon-sun system needed to produce the phases of the moon as seen from earth. However, the treatment used in the present study followed Trundle and Bell (2010) in that *Starry Night*© planetarium software is used for each treatment instead of gathering data using the natural sky. The instruction consisted of four main activities:

1) Students made observations of the moon using *Starry Night*© software for three weeks, sketched these observations and wrote daily reflections.
2) Students analyzed their data looking for patterns in shape, sequence, and direction of their moon observations.
3) Students made mental and physical models of their lunar phase concepts using drawings and 3-D models.
4) Students used their models to predict the dates of subsequent new and full moons.

The male, Caucasian instructor for both treatments had over a decade of experience teaching physics and astronomy at both the high school and college level. Additionally, he had taught the course for three full years that was investigated in the present study. His undergraduate degree is in biology and he holds a Ph.D. in Astronomy. Additionally, he consistently received exceptional teaching evaluations from both students and peers. He previously used inquiry-based activities in his courses based on works such as *Physics by Inquiry* (McDermott, 1996) and was willing to adopt the treatments used in this study. He was interested in student outcomes and is interested in incorporating the findings of the present study into his own teaching in the future.

For this instructional protocol, the instructor only facilitated data collection and guided the students through the learning process. Therefore, the instructor acted as a facilitator. When students asked conceptual questions, the instructor only showed the data, and did not explain the overall concepts. If students had questions, the instructor said they should be included in their daily reflections. The instructor avoided didactic lecturing and the treatments were observed by the researcher to insure this protocol is followed. This implementation fidelity was insured by gently speaking to the instructor in private if there were any diversions from the prescribed instructional protocol.

Over the course of two semesters, four sections of GSCI-162 received two different treatments with the use of *Starry Night*© software each semester. Treatment A and Treatment B were used for one section during each semester. Treatment A consisted of the instructor using *Starry Night*© as a demonstration using the computer projector. Students took data as a class and then interpreted their data in groups of two students on the last day of instruction. The instructor also used *Starry Night*© to test models subsequently developed by the students. Treatment B consisted of students working in groups of two to gather their own data using laptop computers with *Starry Night*© loaded on them. The participants in Treatment B were first taught the basics of operating *Starry Night*© software including simulating sunsets, and facing in the various cardinal points of east, west, north, and south. Operations related to showing the position of the moon were also emphasized. The instructor modeled best practices with the software to help to insure the student participants were not overwhelmed with using it. The students gathered the same data as in treatment A, but they operated the software themselves. Moreover, for treatment B individuals in the groups of two alternated their roles as data takers to insure that both received equal treatments.

For both treatments students worked in groups of two and analyzed their data, built mental and physical models with the use of the Styrofoam® ball (scientific model) activity, and then used their models to make predictions. These predictions were tested using *Starry Night*© software. A summary of the treatments is listed in Table 3 below and a detailed description of the treatments follows.

Detailed Description of Treatments

Participants kept a record of their observations by recording the shape, direction in the sky, and the date and time of the observation using the data sheet (APPENDIX B, Figure 3) for both treatments, A and B. Beyond treatment A, which made use of whole-class gathering of data, and for treatment B where data were gathered in pairs, both treatments were equivalent. Data were gathered only up to the current date for each class meeting. Each group made observations during the waxing phases, from new moon to full moon and then the waning phases up to the 3rd quarter moon. In addition to the daily sketches that students made, they also wrote a daily reflection for each day they make observations. These student reflections included the following elements:

- Drawings made during instruction
- Description of the data gathered
- Possible explanation of the data gathered
- Predictions related to the cause of the lunar phases

- Questions for next class
- Any major surprises from the day

Table 3
Summary of instructional activities

Treat-ment	*Starry Night©* User	Stage of Lesson	Class Session	Classroom Activities
A	Instructor	Data Collection	Days 1-5	Instructor uses *Starry Night©* software in demonstration mode to simulate lunar phases from new to last quarter moon. Students record data on their data sheets.
B	Students	Data Collection	Days 1-5	Pairs of students use *Starry Night©* software to simulate lunar phases from new to last quarter moon. Students record data on their data sheets.
A&B	N/A	Data Analysis *(conceptual change activity)*	Day 6	Students work in pairs and compare their data sheets and develop a moon observation summary chart.
A&B	N/A	Modeling *(conceptual change activity)*	Day 6	Students develop mental, graphic (2-D), as well as 3-D models of the earth-moon-sun system to explain lunar phases as observed from earth.
A&B	N/A	Prediction *(conceptual change activity)*	Day 6	Students predict the shapes, sequences, and positions of the moon for last quarter to new moon, as well as subsequent lunar cycles and eclipses.
A	Instructor	Model Validation *(conceptual change activity)*	Day 6	Instructor uses *Starry Night©* software in demonstration mode to test the predictions that the students made using their models.
B	Student	Model Validation *(conceptual change activity)*	Day 6	Students use *Starry Night©* software to test the predictions that they made using their models.

After the observations were completed, the participants devoted one class period (1.5 hours) to the analysis of their observations. This was a critical part of the treatment which included five separate sections that addressed the following topics: 1) described patterns in the observations such as shape and sequence, 2) applied new concepts and terminology, 3) made a prediction of the length of the overall lunar cycle as well as how the shape changes in a single day, 4) made three-dimensional models to simulate the positions of the moon at various phases and to predict the rising and setting time of the moon at various phases, 5) used the three-dimensional models to simulate the positions of lunar and solar eclipses to directly address the issue of role of shadows in observing the moon. The goal of this treatment was to have students apply their knowledge of lunar phases to construct physical models that show how lunar phases, thus they are applying their knowledge. Thus, the students are using *Starry Night©* in order to construct their own knowledge by use of a constructivist approach to science education, namely by making use of conceptual change. Ideally the students were able to confront their alternative conceptions and replace them with mental models that are scientifically accurate. This was an application of the conceptual change model of science education.

To begin their data analysis, students developed a "Moon Observation Summary Chart", as shown below (APPENDIX B, Figure 4). The chart included not only a description of shape of the moon, but also position in the sky and days observed. Students used their charts to search for patterns in the data, such as changing shapes as well as the changing location of the moon in the sky. Students made sketches of the changing shapes of the moon on their data sheets to illustrate the patterns that they have observed.

Students then used a Styrofoam™ ball (three-dimensional scientific model) and light source, along with their head to represent the earth, to develop a model of the changing relative positions of moon with respect to the Earth and Sun from phases new to 3rd quarter. They used the angular separations measured from the moon to the sun (light bulb) to show how much the moon moves each day. They were able to see that as the lit portion of the ball comes into view, it simulates the various shapes of the lunar phases as the angle changes. The students then simulated the positions of the moon as observed using *Starry Night©* from new to 3rd quarter moon. They then made a sketch in their student reflections indicating how the moon shifts position from new to 3rd quarter moon, keeping the position of the sun and earth constant. They also developed an explanation of the cause of lunar phases and describe it in their student reflections.   Terminology such as the names of the various shapes was introduced so that the participants will have a uniform set of descriptors for the shapes. The terms of waxing and waning phases were also introduced.

At the end of the treatments the instructor facilitated a discussion that reviewed the four key concepts associated with the scientific explanation of lunar phases. This was to insure that the students themselves construct their own mental models during instruction. These concepts include:

- Half of the moon is illuminated by the sun
- The portion of the illuminated half seen from Earth varies over time
- The relative positions of the earth, sun, and moon determine the portion of the lighted half seen from Earth
- The moon orbits Earth

The students made a prediction of how the lunar phases will change once the Styrofoam ball representing the moon passes the position for 3rd quarter moon. They made sketches in their reflections describing their predictions along with an explanation of their model on which they have based their predictions. They then tested their predictions by using the 3-D models as well as testing their predictions using *Starry Night©*. They were able to revise their new mental models of the cause of the moon phases in accordance with conceptual change. They reported in their reflections discussing how their predictions performed and any discrepancies they encountered. They then sketched a revised model based on what they learned from testing their mental model of lunar phases.

The students then worked in groups on a series of questions to extend their knowledge related to the time of day certain lunar phases are visible and location in the sky. These extension questions are listed below:

i. Are the sun and the moon ever in the sky at the same time?
ii. Where is the sun relative to the moon when the moon is full?
iii. Where is the sun relative to the moon when the moon is first quarter?
iv. Where is the sun relative to the moon when the moon is third quarter?

Instrumentation and Data Collection Methods

The LPCI was administered both before and after instruction to each participant to measure their scientific understanding of lunar phases. The LPCI (see Appendix A) is a twenty-question forced-choice instrument with nine additional demographic and attitudinal questions. The structured interviews are based on the interview protocol for lunar phase concepts described in Trundle, Atwood, and Christopher, (2002, 2006, 2007a,b) and Bell and Trundle (2007, 2008). Responses to the LPCI were used to purposefully sample six students from each class. There were two randomly selected students from the low, mid, and high scores for the pre-instruction results. Each sampled student was interviewed both before and after instruction, for a total of 48 interviews over two semesters. The purpose of the interview was to be able to ask students probing questions as they discuss lunar phases to more completely assess their conceptual understanding of the phenomenon of lunar phases. Several follow-up questions were added to the interview protocol for the post-instruction interview in order to obtain a more detailed understanding of the student opinions of the treatments and to investigate the principle components and relative effectiveness of these components in the learning process. The interview protocol was as follows:

Task 1. Each participant was asked to draw the shapes they expect to see, or those they had seen, depending on whether it is the pre or the post instruction interview. The participants were then asked to explain how they drew the shapes so as to minimize misinterpretation of the shapes by the researcher.

Task 2.   Participants were asked to predict if the moon shapes change in a predictable pattern (pre) or describe their changes (post).

Task 3. For both pre and post interviews, the participants explained what had caused the moon shapes.

Task 4. For both pre and post interviews, participants used 3-D models of the earth, moon and sun to demonstrate their ideas of lunar phases, the participants also verbally explained their models.

Task 5. For both pre and post interviews, the participants completed a card sorting activity in which they attempted to correctly put the lunar phases in order.

Task 6. For both pre and post, the interviewer probed the participants' responses in order to assure correctly assessing their conceptual understanding during each task.

Follow-up questions for post-instruction interview in addition to Trundle protocol Trundle et al. (2002):

Question 1. List the instructional activities you remember.

Question 2. Did your understanding of lunar phases change during instruction?

Question 3. Which of these activities, if any, impacted your views?

Question 4: Would you use this type of instruction for lunar phases with your future students?

Question 5: (Referring to Question 4) If so, what would you change, if anything?

Question 6: During the three weeks of instruction, did you happen to see the real moon in the sky?

Question 7: (Referring to Question 6) If so, how did it affect your understanding of lunar phases?

Question 8: What role did the daily reflections that you wrote during instruction influence your scientific understanding of lunar phases?

Question 9: How did your views change after instruction compared to before instruction?

The data were analyzed to look for differences in the conceptual gains for the entire group using descriptive statistics by calculating the difference between pre and post instruction for the groups. Both instruments were used to calculate the students' level of scientific understanding of lunar phases. The descriptive and predictive statistics were used to determine if the increase in scientific understanding is significant and to answer the following research question:

1. How do pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases change after inquiry-based, conceptual change instructional intervention based on a computer simulation?

The effectiveness of each of the two treatments were measured by both instruments and will be compared to see if there is a statistically significant difference between the two treatments using a statistical technique called analysis of variance (ANOVA). Therefore, the two instruments were important in answering the following research question which is:

2. Are there differences in achievement between pre-service elementary teachers who collect moon observation data in a whole-class setting versus those who collect this data working in pairs with the computer simulation?

Additional interview question were asked during the semi-structured interview after instruction. These questions probed to determine the key components during instruction for the students. The quality of the student reflections were also used to answer the research question:

3. How do pre-service elementary teachers perceive the relative value of each component of the instructional intervention as related to their attainment of scientific conceptions of moon phases?

The LPCI was administered to all 69 participants and the semi-structured interview of the Trundle protocol was administered to a purposely selected sampling of students. The sampling was performed by ranking each of the four sections studied by LPCI score and within those sections assigning LPCI scores to a top, middle, or bottom group. Two participants from each group were then randomly selected to participate in the interview before and after instruction. This data from the LPCI and semi-structured interview was used to answer the last research question:

4. How does the forced-choice LPCI instrument compare with the semi-structured Trundle protocol in measuring the conceptual understanding of lunar phases among pre-service elementary teachers?

The treatments for each of the groups were observed and notes taken in order to record the actions of the instructor and students. This was done to confirm the prescribed treatments are indeed being used. Observations of the classroom as well as transcripts from the treatments were analyzed for any deviations from the prescribed treatments. These deviations were used to determine if both treatments were equivalent, other than the planned differences in students gathering data. Guidelines for the classroom observations were as follows:

- Instructor provides adequate introduction to the use of *Starry Night*© software
- Students understand the software and are able to make sketches from the display
- Instructor leads the students in their inquiry and does not didactically teach.
- Students are engaged and are taking observations
- Students work in pairs to develop their mental models of lunar phases.
- The instructor-student interaction facilitates inquiry-based instruction.

Reliability and Validity of Instruments

Inter-rater reliability is a measure of how well do different raters agree with their coding of the same interviews. Miles and Huberman (1994) state that this type of reliability should be approach ninety percent, which is calculated using the ratio of number of agreements divided by the number of ratings times one hundred percent. The inter-rater reliability for the structured interview protocol was established in previous studies (Trundle et al., 2002, 2006, 2007a,b) and was consistently greater than ninety percent. The current researcher (William Alexander) scored a 100% agreement with another previously trained researcher with a training set of interviews in May, 2003. The reliability for the LPCI was established by Lindell and Olsen (2002) and found the *Cronbach's alpha* for the pretest to be 0.55 and for the posttest to be 0.75. The pretest is lower than the generally acceptable value of 0.70. Lindell and Olsen (2002) indicate that this lower value could be due to student guessing. Therefore, the results from the LPCI should be carefully considered especially for the pretest results because the pre-test LPCI results may tend to have lower reliability. The reliability of the Abridged-LPCI is assumed to be equivalent to the reliability of the LPCI since the Abridged-LPCI was developed using a purposeful sampling of LPCI questions that directly aligned with the content covered in the Trundle interview protocol.

Internal and External Validity

Miles and Huberman (1994), state that internal and external validity are an indication of the findings being deemed valid and credible and are not easily quantifiable. The internal validity rests in the results being attributed to a specific variable that is manipulated. It is important that as many variables as possible to be controlled, so that any differences between the treatments can be attributed to the independent variable manipulated by the researcher and not extraneous variables.

According to Shadish et al. (2002), some possible threats to internal validity could include: Ambiguous variable designation, participant selection, history during treatment, maturation of students' understanding, regression of extreme scores, attrition of participants during the treatment, testing effect in which the instrument also contributes to the treatment, the instrument may change during treatment so as to mimic being part of the treatment, and finally the interaction of two or more of the above threats to validity. For this research, the variables were well-defined because one classroom received treatment A and the other classroom received treatment B. The participants were students who were registered for two regularly scheduled sections of a class that provides basic space science content for students preparing to be preK-8 teachers. These classes were chosen for this research because of their convenience as well as access to the specific instructor. These classes were also chosen because they are similar classes studied by Trundle et al. (2002, 2006, 2007a,b) and Bell and Trundle, (2006, 2007, 2008) because they are all groups of undergraduates preparing to be teachers. The treatments (A & B) were be randomly assigned to the two sections. The make-up of the classes is by chance and not assigned. Both sections were on Monday and Wednesday afternoons. One section was at 12:20-2:15 PM and the other section was at 2:30-4:25 PM. Any differences between the two sections were measured in the pre-treatment assessment of the participants and were taken into consideration when calculating conceptual gain after instruction.

Nothing significant related to lunar phases happened outside of class for these participants during the treatment, since the treatment only lasted for three weeks. During the treatment, there were no efforts to encourage students to make observations of the lunar phases using the real sky. All participants obtained their lunar observations via *Starry Night*© software. Therefore, the history during the treatment was not a threat to internal validity. The maturation of the students during the treatment was not a threat to internal validity, since the treatment was relatively short at only three weeks and that no other courses they were taking covered topics related to lunar phases. The regression f scores was not a problem, since the most extreme scores were not studied alone. The means of the groups will be compared to gains in conceptual understanding for the groups as a whole, not an extreme subset.

Attrition, or the loss of participants during the study, could very well have been a threat to internal validity, however all those enrolled in the classes agreed to participate in the study. Each participant completed the study. However, participants were allowed to leave the study at any time. Full participation was aided with offering incentives of nominal value to participants. Also, only participants that had both a pre and post score were used to compute the pre and post instruction means for each treatment. This was not an issue, since there was full participation. The testing effect has been shown not to be a factor in such studies as shown in Trundel et al. (2002). Trundel et al. (2002) used a control group that measured conceptual understanding of lunar phases only at post-instruction and the results for the control were comparable to those treatment groups that received the pre-instruction assessment. The semi-structured interview protocol did not change during the treatment, so the pre and post results measured the same variable, which is the participant's level of scientific understanding of lunar phases.

The validity of the LPCI was established as described in Lindell and Olsen (2002). Prior to field testing, the content was examined by subject matter experts. Item analysis was performed on each question that included concentration and discrimination analysis. Lindell and Olsen (2002) state that in order for an item to be included in the LPCI, the discrimination values must be greater than 30% because this value indicates that a specific item can discriminate between high and low overall scores. Additionally, Lindell and Olsen (2002) state that pre and post concentration factors for an individual item should be greater than 20%, or slightly more than concentrated on the correct response than one would expect for equal distribution across all answers. The validity of the Abridged-LPCI is assumed to be equivalent to the validity of the LPCI since the Abridged-LPCI was developed using a purposeful sampling of LPCI questions that directly aligned with the content covered in the Trundle interview protocol.

The classroom instruction was videotaped and later reviewed to help to insure that both sections received equivalent instruction, other than the manipulated variable. To insure internal validity, there were multiple sources of data considered. The interviews, LPCI, as well as student reflections were independent assessment instruments that improve internal validity through triangulation.

The external validity is determined by if the study is able to be generalized to other situations. Since the students participating in this study were all preparing to be teachers, the argument for their applicability to other similar institutions with teacher preparations can be made. A key to this transferability is that the study is conducted with pre-service teachers which is the same general group used in previous studies.

Data Analysis

The basis for the data analysis is the use of systematic data analysis as described by Miles and Huberman (1994). This technique used predetermined codes to categorize the data that can then be used with statistical analysis techniques.

The scientific concepts covered in the semi-structured interview developed in Trundle et al. (2002) include:

1. students correctly describing and drawing shapes of lunar phases
2. students correctly sequencing the lunar phases
3. students correctly being able to explain the four scientific causes for lunar phases

The scientific causes of the lunar phases include:

a. Half of the moon is illuminated by the sun
b. The portion of the illuminated half seen from Earth varies over time
c. The relative positions of the earth, sun, and moon determine the portion of the lighted half seen from Earth
d. The moon orbits Earth

The scientific lunar phase concept domains covered in the LPCI and described in Lindell and Olsen (2002) are given as:

1. Period of the Moon's orbit around the Earth
2. Direction of the Moon's orbit around the Earth as viewed from a point above the north pole
3. Period of the Moon's cycle of phases
4. Motion of the moon
5. Phase and Sun-Earth-Moon positions
6. Phase-location in sky-time of observation relationship
7. Cause of Lunar Phases
8. Effect on lunar phase with change in location on Earth

For the semi-structured interview, the data for moon shapes was analyzed for accuracy of the shapes such as full, crescent, gibbous and new. Specific attention was paid for instances of a false gibbous where the shape looks more like the initial stages of a lunar eclipse. The drawings of the crescent shapes should be made to begin at the poles of the moon, so as to not make the crescents too small. The sequence of shapes should show a gradual progression, starting with either the new moon then cycling through the waxing and waning phases. Alternatively, the sequence may begin with the full moon and then cycle through the waning and waxing phases since the full moon can also be used as the starting point of the lunar cycle as well. The data analysis approach for levels of conceptual understanding for lunar phases was developed by Trundle and various collaborators (Bell et al., 2007; Bell & Trundle, 2007; and Trundle et al., 2002, 2004, 2006). The coding shown in table 4 and the types of scientific understanding were developed by Trundle and collaborators using previous research (Callison & Wright, 1993, Stahly et al., 1999) to develop an understanding of the various alternative conceptions that participants might have. The coding of each interview adapted the protocol as described in Trundle et al. (2002, 2007) and is described and listed in Table 4 below.

Table 4
Interview Coding Guide

Code	Meaning of Code
SciOrb	Moon orbits earth.
SciHaf	Half of moon is illuminated, that half toward sun.
SciSee	Part of the illuminated half we see determines phase.
SciEMS	Relative positions of earth, sun, and moon determine the part we see.
AltEcl	Dark part of the moon in earth's shadow; phases caused by earth's shadow.
AltRot	Earth's rotation on axis causes phases.
AltHel	Heliocentric: Moon and Earth do not orbit Sun or Sun between Earth and Moon.
AltETilt	Tilt of the earth on its axis causes moon phases.
AltGeo	Moon phases due to a viewer's geographic position on the earth.
AltClo	Clouds or weather conditions causes phases.
AltFrg	More than one above Alt given (list them).
AltOth	Reason other than any of the above given. Provide description and add to coding system.
NoCU	No conceptual understanding evident or no response given.

For each interview, there was a numeric assignment for each of the categories of conceptual understanding in order to perform statistical calculations. The numeric values ranged from zero for no conceptual understanding to 10 points for an interview were where the participant exhibits a full understanding of the cause of the lunar phases including all four scientific elements with no alternative conceptions. The number of alternative conceptions held by a participant, even who expressed all four scientific elements, resulted in a lower assigned score. Therefore, there were intermediate conceptual understanding levels as shown in APPENDIX C. This scoring rubric is based on Trundle et al. (2010).

These numeric values were then be used in an analysis of variance (ANOVA) to measure if the differences between the mean scores after instruction are statistically significant and what they can be attributed to. The ANOVA also controlled for any differences in the two groups prior to instruction. The same procedure was also performed with the LPCI results.

The results were also be compared to other studies to investigate whether obtaining data for only the new to last quarter phases of a lunar cycle is enough to obtain similar gains when compared to other studies which used more days of observations. Analysis of variance (ANOVA) was then used to measure differences between means for each of the treatments. These analyses made it possible to answer the following research questions:

1. How do pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases change after inquiry-based, conceptual change instructional intervention based on a computer simulation?
2. Are there differences in achievement between pre-service elementary teachers who collect moon observation data in a whole-class setting versus those who collect this data working in pairs with the computer simulation?

In the process of analyzing the LPCI results and comparing them to the sem-structured interview results, an Abridged-LPCI instrument was developed that focused more directly on the concepts addressed in the interview. Those items in the LPCI that did not directly match a concept in the interviews were removed to establish the Abridged-LPCI. The relative performance of the LPCI and the Abridged-LPCI were compared in order to determine if there might be a quicker alternative to the time-consuming interview protocol.

Additional data for this study also recorded in student observational reflections and students gathered lunar observations using *Starry Night*© for the period of starting with the new moon and ending at the last (3rd) quarter moon approximately 21 days later. Cropp (1980) showed that effective student journals in a Physical Geology class contained a significant number of observations and student-generated questions. Hyers (2001) showed that effective student journals for an introductory earth science class included observations, reactions, as well as development of thought.

Since the reflections that students wrote during instruction provide a rich source of data, they were analyzed for the following criteria:

- The number of words in each student reflection
- The number of questions the student generates in each reflection
- The number of student predictions in each reflection

Additionally, the quality the engagement of each student reflection was made on a scale of 0 to 3 with rating of:

1. **No engagement:** It is apparent that the student is not an active participant and little or no insight can be gained regarding their understanding of lunar phase concepts. The reflection is incoherent and/or not understandable (e.g., incomplete sentences).
2. **Low engagement:** The student reflection contains segments of conceptual understanding of lunar phases, but they are not connected in any way (e. g. student uses complete sentences but no connections are made).
3. **Adequate engagement:** The student is an active participant and it is obvious they are participating in the ongoing assignment. The reflection is understandable but lacks a comprehensive model of

lunar phases (e.g., student uses complete sentences, some connections are made, but does not explicitly state a model).

4. **High engagement:** The student is an active participant and they are participating in the assignment. The reflection is understandable and provides a comprehensive model of lunar phases (e.g., student uses complete sentences, connections are made, and student explicitly states a model).

The engagement ratings did not measure the scientific accuracy of the lunar phase conceptual models students made, but simply their engagement in the activity as measured by the student reflections. There were four scores associated with a student's level of engagement which included: number of words, number of questions, the number of predictions, and the qualitative engagement score will be determined by the above rubric.

These numeric values were then used in an analysis of variance (ANOVA) to measure if the differences between the mean scores after instruction can be attributed to the student reflection results. The ANOVA also controlled for any differences in conceptual understanding between the two groups prior to instruction. The same procedure was also performed with the LPCI results as the dependent variable to test if the LPCI results have an effect on the conceptual gains.

The results from the analysis of the student reflections were combined with additional questions during the student post-instruction interview that probed the students' feelings about the most important part of instruction. There was a ranking of factors related to instruction to answer the following research question:

3. How do pre-service elementary teachers perceive the relative value of each component of the instructional intervention as related to their attainment of scientific conceptions of moon phases?

The concept domains of the LPCI were matched with the appropriate elements of the semi-structured interviews. A concept map was developed which helped to graphically represent similar and dissimilar concepts. Pearson correlations of the LPCI and interview data were also calculated in order to determine if there are any statistically significant correlations between the two instruments. This comparison of the LPCI and semi-structured interview was used to answer the last research question:

4. How does the forced-choice LPCI instrument compare with the semi-structured Trundle protocol in measuring the conceptual understanding of lunar phases among pre-service elementary teachers?

Unit of Analysis for the ANOVA

The individual student was used as the unit of analysis for all ANOVAs. This approach is contrary to the most conservative procedures for science education research, which would prescribe each class as the unit of analysis, since random assignment was conducted on the class level as is described in Stoker et al. (1981). This type of assignment distributes any possible non-experimental factors across both treatment groups evenly, limits threats to internal validity, and also increases arguments for causation. However, Shadish et al. (2002) state that when experimental units, such as students, cannot be randomly assigned, a Quasi-experimental design can be used, as long as special attention is given to threats to internal validity.

To this end, the treatments were randomly assigned to each of the classes and there is no reason to think that students registered for the sections in a non-random fashion. The students were allowed to freely choose between equivalent sections that met on the same days of the week, in the afternoons, separated by two hours.   Additionally, the pre-test results were statistically the same for each of the treatment groups and treatments A and B were identical other than treatment A as a whole-class and treatment B where students used *Starry Night©* in pairs as shown in Table 5. Further, after investigating each treatment group for parameters such as age, gender, college major, ethnicity, there is nothing about the treatment groups that would suggest that they are not homogeneous. Moreover, all four of the sections had the same instructor which also reinforces that the students received the same treatment, other than "whole-class versus pairs" part of the treatments. These parameters suggest that individual students were randomly enrolled into the two treatments, at least in regard to constructs of interest. Therefore, the student as the unit of analysis was used. The design of this study is classified as quasi-experimental, as opposed to an experimental design.

Researcher as Instrument

I see myself as a researcher with a strong interest in astronomy and find the topic of phases of the moon as being a critical issue in science education. My training includes a BS degree in Physics and a MS degree in Physical Science. Professionally I served as an assistant professor at James Madison University in Harrisonburg, VA from 2001 to 2011. Additionally I have served as the director of the John C. Wells planetarium at JMU from 2007-2011. I am currently a patent examiner with the US Patent & Trademark

Office in Alexandria, VA. My personal aspirations are to provide research-based science education curriculum approaches to the Patent Training Academy.

I consider myself as a post-positivist researcher and feel that a constructivist approach holds the most promise for teaching astronomy concepts because students are allowed to confront their alternative conceptions. I feel as though many researchers have rushed into qualitative research studies and instruments such as the ADT and the LPCI not only do they have problems with measuring conceptual but they also are unable to help the researchers understand anything that happens during the learning process. I feel as though I am biased toward qualitative studies that use interview techniques and that I am skeptical of studies that use the ADT, the LPCI, and similar instruments.

I feel as though using planetarium software, such as *Starry Night*© has greatly improved my teaching and I think it could also be equally important to other instructors. Personally I feel as though students should be encouraged to view the night sky, I realize that it is difficult to use the sky as a natural laboratory. From the literature I feel as though the advantages of making observations in the classroom using *Starry Night*© outweigh the advantages of making observations outdoors.

While at JMU, I taught sections of the GSCI-162, Science of the Planets, which are the courses I am investigating in this study. I distanced myself from teaching these courses since 2008 by not teaching a section of GSCI-162. I am familiar with the instructor for the courses, Geary Albright, but have never team-taught with him and I had no supervisory relationship with him. We were simply colleagues.

CHAPTER 4: FINDINGS OF THE STUDY

General Description of Data

The data for this study were obtained over two semesters with two classes studied each semester for a total of four sections and 69 participants. Each participating student completed an abridged version of the LPCI before and after instruction. The LPCI was abridged to better represent the content assessed in the Trundle interview protocol Trundle et al. (2002, 2007). The gains for the LPCI as well as the Abridged-LPCI were calculated. The daily reflections written by each of the students were also analyzed for this investigation. The reflections provided the following measures: word count, number of questions posed, number of predictions made and the level of engagement as determined by rubric. The reflections were also a source of frequent comments from the participants. The frequencies of common comments from the participants were recorded as well as common misrepresentations of the lunar shapes drawn by the participants. A purposefully selected group of 6 participants from each class section, for a total of 24, were assessed using the Trundle interview protocol both before and after treatments.

Descriptive Statistics

LPCI and Abridged-LPCI Results

Below (Table 5) is given an overview of the descriptive statistics by using the instruments of LPCI, Abridged-LPCI, student reflection, as well as interview. Both LPCI and Abridged-LPCI indicated an increase in conceptual understanding of lunar phases when post-instruction results are compared to pre-instruction results. The effect size for the LPCI results for comparing post-instruction to pre-instruction results in 1.0, which can be categorized as a large effect size according to Fan (2001). Similarly the effect size for the Abridged-LPCI was found to be 1.1, which can also be categorized as “large.” In terms of effect size calculations, both forced-choice instruments were similar. The mean of LPCI Treatment-A for pre-instruction (M=36.6, SD=16.5) is not statistically different than LPCI Treatment-B for pre-instruction (M=36.8, SD=16.3) using a two-tailed, unpaired t-test, $t(67) = 0.06$, $p=0.95$. Similarly, The mean of Abridged-LPCI Treatment-A for pre-instruction (M=41.4, SD=20.2) is not statistically different than Abridged-LPCI Treatment-B for pre-instruction (M=40.8, SD=20.2) using a two-tailed, unpaired t-test, $t(67) = 0.13$, $p=0.89$.

Comparison of LPCI and Abridged-LPCI

In order for the LPCI to better match the conceptual understandings measured during the interview, an Abridged-LPCI was developed. This Abridged-LPCI differed from the Abridged-LPCI developed in Bell et al. (2007) because it used only existing questions from the original LPCI and did not re-write or add additional questions. Questions from the original LPCI that did not correspond to the semi-structured interview protocol were simply removed. The initial results from the comparison of the two instruments are given below in Figure 1. Note that there is a moderate amount of overlap between the two instruments. More topics are covered in the LPCI than in the semi-structured interview protocol. Therefore, to compare the interview to the LPCI, some questions must be eliminated from the LPCI as shown in Appendix A. Any question related to the categories: Period of the moon’s orbit, direction of the moon’s orbit, or

Table 5
Conceptual Understanding Measured by Interview, LPCI, and Abridged-LPCI

Measurement	*n*	Mean (SD) Pre-Instr.	Mean (SD) Post-Instr.	Gain (SD)
Conceptual Understanding from Interview (0-10 pts.)	24	4.5 (1.8)	8.1 (2.4)†	3.6 (2.0)
LPCI (0-100 pts.)	69	36.7 (16.4)	52.9 (15.3)†	16.1 (15.3)
Abridged-LPCI (0-100 pts.)	69	41.1 (20.0)	64.4 (20.7)†	23.3 (20.7)
LPCI Treatment A (whole class)	34	36.6 (16.5)*	55.1 (15.2)†	
LPCI Treatment B (pairs)	35	36.8 (16.3)*	50.6 (15.2)†	
Abridged-LPCI Treatment A (whole class)	34	41.4 (20.2)**	65.8 (20.8)†	
Abridged-LPCI Treatment B (pairs)	35	40.8 (20.2)**	63.2 (20.8)†	

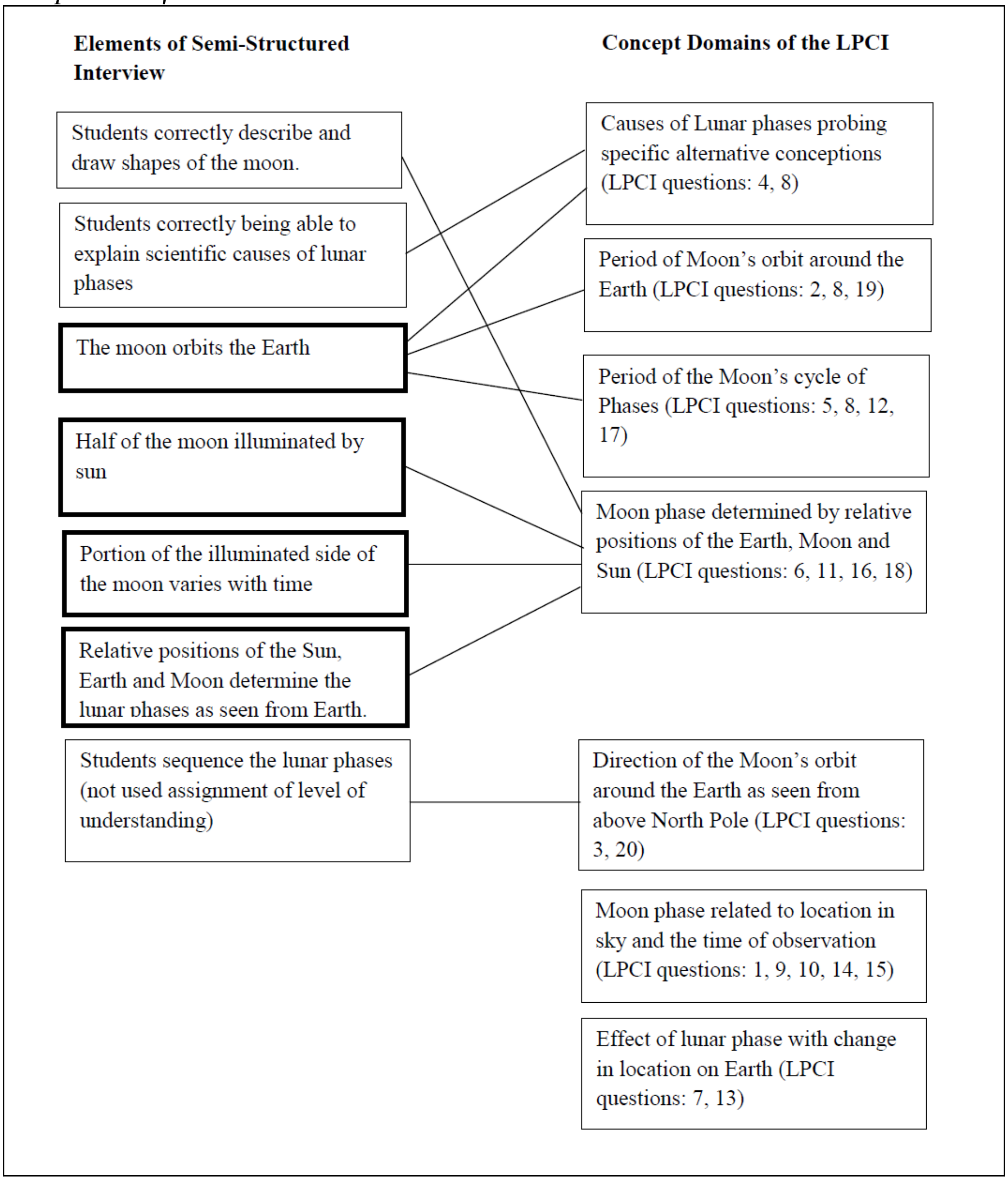
*Indicates that the two means are not statistically different.
**Indicates that the two means are not statistically different.
†Indicates a large effect size when comparing pre and post instruction means.

effect of lunar phase with location on the earth will be eliminated. The questions from the LPCI that meet these criteria and should be eliminated are: 1, 3, 7, 9, 10, 13, 14, 15, and 20. This represents 9 out of 20 questions from the LPCI being eliminated in order to better mimic the results from the semi-structured interview protocol which results in a 55% overlap between the two instruments.

The LPCI goes deeper into conceptual understandings of lunar phases by addressing time of day and position of an observer on Earth, while the semi-structured Trundle protocol addresses minimums of conceptual understandings for pre-service teachers that will be teaching lunar phase concepts in the future.

Figure 1

Comparison of semi-structured interview to the LPCI



Interviews

The conceptual understanding increased for the 24 participants as a whole. The interviews were coded according to rubric as shown in Table 4 in Chapter 3. The codes were developed from the scoring rubric for the interview protocol (Trundle et al., 2010). In general, the number of students holding a scientific understanding of the cause of lunar phases increased and the number of students holding alternative conceptions decreased when post-instruction results are compared to pre-instruction results. The one exception is that the alternative conception "AltETilt" had one student holding that view in pre-instruction and one student holding that view post-instruction. The most prevalent alternative conception prior to instruction is that the Earth's shadow causes the lunar phases (AltEcl) and is also the most prevalent alternative conception after instruction. The second most prevalent alternative conception prior to instruction is that the moon orbits the Sun and not the Earth (AltHel), however it differs from the most prevalent alternative conception in that none of the participants held this view after instruction.

Table 6 includes three alternatives that have not been previously reported. They include: Alternative Orbital Speed, Alternative No Orbit, and Alternative Large Sun. These newly identified alternative conceptions were not expressed by participants after instruction.

Table 6

Frequencies of interviewed participants' conceptual understandings of the cause of moon phases

Type of conceptual Understanding	Criteria and [Codes]	Freq. Pre-Instruction *(n=24)*	Freq. Post-instruction *(n=24)*
Scientific	All four scientific criteria included:	4(17%)	17(71%)
	• Half the moon illuminated by sun [SciHaf]	7(29%)	17(71%)
	• The portion of the illuminated half Seen from earth varies with time [SciSee]	7(29%)	22(92%)
	• The moon orbits the earth [SciOrb]	21(88%)	24(100%)
	• The relative positions of the earth, Sun and moon determine the portion Of the lighted half seen from earth [SciEMS]	21(88%)	23(96%)
Alternative Eclipse	Earth's shadow on the moon causes moon phases (eclipse) [AltEcl]	16(67%)	8(33%)*
Alternative Heliocentric	Moon orbits the sun but not the earth causing moon phases [AltHel]	3(13%)	0(0%)
Alternative Rotation	Earth's rotation on its axis causes moon phases [AltRot]	2(8%)	0(0%)
Alternative Earth's Tilt	Earth's tilt toward or away from the the moon causes moon phases [AltETilt]	1(4%)	1(4%)**
Alternative Clouds	Cloud cover causes moon phases [AltClo]	1(4%)	0(0%)
Alternative Geographic Position	Moon's position relative to an observer's position on the earth determines moon phase [AltGeo]	1(4%)	0(0%)
Alternative Orbit Speed***	Moon orbits earth in 24 hours causing the moon phases	1(4%)	0(0%)
Alternative No Orbit***	The moon is in the same position for all moon phases	1(4%)	0(0%)
Alternative Large Sun***	Sun is larger than earth so light goes light goes around earth causing the Full moon	1(4%)	0(0%)

*4 of the 8 also correctly described all four scientific criteria post-instruction.
** Participant also correctly described all four scientific criteria post-instruction.
*** Alternative conceptions identified in this study and not previously reported

Lunar Phase Sequencing during Interview

Lunar Phase Sequencing during Interview. The student sequencing activity during pre and post instruction interview are given in Table 7. The number of participants that sorted and sequenced the lunar phases incorrectly decreased when comparing post-instruction to pre-instruction. Moreover, the number of participants that incorrectly sequenced the lunar phases by simply reversing them also decreased when comparing post-instruction to pre-instruction. Conversely, the number of students that correctly sorted and sequenced the lunar phases increased when comparing post-instruction results to pre-instruction results. These results indicate that some students were replacing their alternative conception that the lunar phases are random with the scientific conception that changes in lunar phases are gradual and follow a predictable pattern.

Table 7
Sequencing of Lunar Phase Drawings Pre/Post-Instruction from Interview: Outcome, sample size, number giving response, percentage

Outcome	*n*	Number (%) Pre-Instr.	Number (%) Post-Instr.
Sequencing Incorrect	24	7(29.2%)	5(20.8%)
Sequencing Incorrect (reversed)	24	10(41.7%)	9(37.5%)
Sequencing Correct	24	7(29.2%)	10(41.7%)

Interview Results. Table 5 shows the interview results indicate increases in conceptual understanding when post-instruction results are compared to pre-instruction results. The effect size was calculated to be 1.7, which can be categorized as “large” according to Fan (2001). These results were similar to the LPCI and Abridged-LPCI results that exhibited similar increases as also shown in Table 5.

Reflections

All of the participants completed their daily student reflections, for a sample size of 69. The reflections were not scored for accuracy of scientific concepts. The engagement of the participants was the purpose of the students keeping their daily reflections. The student reflections (Table 8) showed a great variation in the word count with a mean of approximately 300 with a standard deviation of approximately 143. The mean number of questions students developed in their reflections was much less than one per student, while the number of questions developed by each student was greater than one, but less than two. The mean level of engagement for the reflections resulted in an overall "adequate" engagement. The level of engagement was determined with the engagement rubric described in Chapter 3.

Table 8
Writing of a Reflection during Instruction

Measurement	*n*	Mean (standard deviation)
Student Reflection Word Count	69	299.2(142.7)
Student Reflection Number of Questions	69	0.1(0.4)
Student Reflection Number of Predictions	69	1.6(1.8)
Student Reflection Level of Engagement	69	2.0(0.7)

LPCI

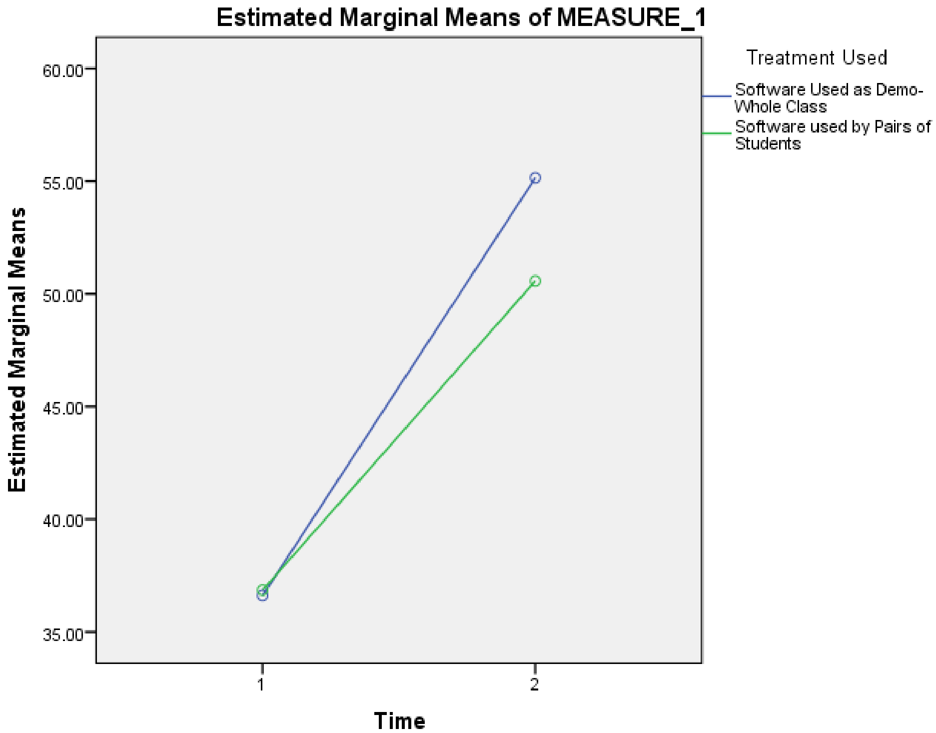
For the analysis of the LPCI scores, a two-way mixed-design analysis of variance (ANOVA) was performed. The between-subjects factor was treatment (whole-class vs. pairs) and the within-subjects factor was time (pretest vs. posttest). A significance level of 0.05 was used for this single-test ANOVA analysis. Table 9 shows the results of the analysis.

Table 9
Analysis of Variance: LPCI Score as a Function of Treatment and Time

Source	*df*	*F*	*p*
Treatment	1	0.418	0.520
Between error	67	–	–
Time	1	77.288	0.000
Treatment × time	1	1.724	0.194
Within error	67	–	–

The time effect was significant, $F(1, 67) = 77.288$, $p \approx 0.000$ as shown in Table 11. The posttest mean ($M = 52.9$) was significantly higher than the pretest mean ($M = 36.7$), which was previously stated to be a “large” effect size. The treatment and time interaction was non-significant, $F(1, 67) = 1.724$, $p = 0.194$. The four cell means are described in Figure 2. Figure 2 shows that the two lines do not significantly cross which graphically supports that there is not a significant interaction of time and treatment.

Figure 2
Mean LPCI scores for the four groups



The analysis shows that the effect of treatment (A vs. B) was non-significant, $F(1, 67) = 0.418$, $p = 0.520$. The mean LPCI score for the demo group ($M = 45.9$, Std. Error = 2.4, SD = 13.9, $n = 34$) and the mean LPCI score for the pair group ($M = 43.7$, Std. Error = 2.4, SD = 15.0, $n = 35$) were not significantly different. The effect size when comparing the demo group and the pair group was calculated to be 0.1, which is a "low" effect size according to Fan (2001). For this analysis, the equality-of-variances assumption was satisfied for the pretest measure, $F(1, 67) = 2.413$, $p = 0.125$, but was not satisfied (violated) for the posttest measure, $F(1, 67) = 6.139$, $p = 0.016$. However, this is considered not to have a significant effect on the results (K. Nashimoto, personal communication, July 19, 2012). This is graphically represented in Figure 2 by the two lines not substantially crossing.

Abridged-LPCI

For the analysis of the Abridged-LPCI scores, a two-way mixed-design analysis of variance was performed. The between-subjects factor was treatment (demo vs. pair) and the within-subjects factor was time (pretest vs. posttest). A significance level of 0.05 was used for this single-test ANOVA analysis. The results are shown in Table 10.

Table 10
Analysis of Variance: Abridged-LPCI Score as a Function of Treatment and Time

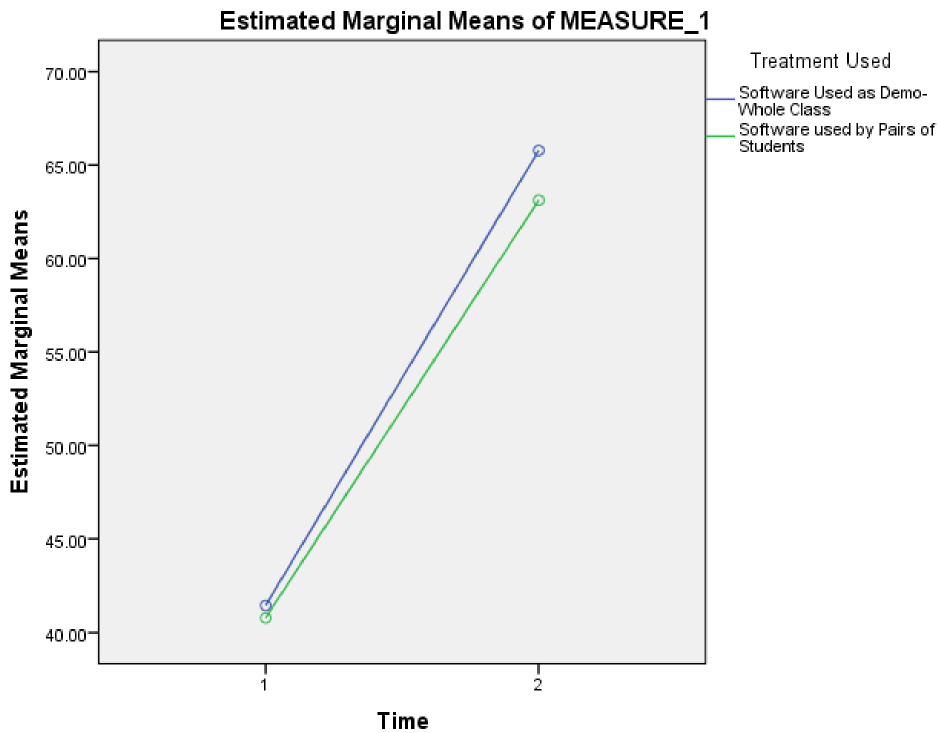
Source	*df*	*F*	*p*
Treatment	1	0.153	0.696
Between error	67	–	–
Time	1	86.481	0.000
Treatment × time	1	0.158	0.692
Within error	67	–	–

The time effect was significant, $F(1, 67) = 86.481$, $p \approx 0.000$ as shown in Table 14. The posttest mean ($M = 64.4$) was significantly higher than the pretest mean ($M = 41.1$), which was previously described as having a "large" effect size. The treatment and time interaction was non-significant, $F(1, 67) = 0.158$, $p = 0.692$. The four cell means are described in Figure 3. Figure 3 shows that the two lines do not significantly cross which graphically supports that there is not a significant interaction of time and treatment.

The analysis shows that the effect of treatment (A & B) was non-significant, $F(1, 67) = 0.153$, $p = 0.696$. The mean Abridged-LPCI score for the demo group ($M = 53.6$, Std. Error = 3.0, SD = 17.6, $n = 34$) and the mean Abridged-LPCI score for the pair group ($M = 51.9$, Std. Error = 3.0, SD = 17.6, $n = 35$) were not significantly different. The effect size when comparing the demo group and the pair group was calculated to be 0.1, which is a "low" effect size according to Fan (2001). For this analysis, the equality-of-variances assumption was satisfied for the pretest measure, $F(1, 67) = 2.177$, $p = 0.145$, and was also satisfied for the posttest measure, $F(1, 67) = 3.519$, $p = 0.065$. Therefore, both satisfy the equality of variance test. This is graphically represented in Figure 3 by the two lines not crossing.

Figure 3

Mean Abridged-LPCI scores for the four groups



Interviews Compared to LPCI and Abridged-LPCI

Figure 1 showed how each question of the LPCI was or was not aligned with the interview questions. The questions that were not aligned with the interview were removed to establish the Abridged-LPCI. As a statistical comparison of the various instruments, correlations were calculated comparing LPCI scores to interview results and correlations were also calculated comparing the Abridged-LPCI scores to the interview results.

Correlations of LPCI Scores to Interview Results

Table 11 shows that the LPCI for the pre-instruction is correlated with the pre-instruction interview and is statistically significant (r = 0.613, p = 0.001) at the alpha = 0.017 level. The modified-alpha level of 0.017 was used after the alpha of 0.05 was modified using the Bonferroni correction for 3 tests so that the adjusted alpha = 0.05/3 = 0.017. The statistically significantly correlation of the pre-instruction interview and pre-instruction LPCI is shown by scatter plot in Figure 4. The correlations for post-instruction and Gains (post-pre) are not statistically significant at the alpha = 0.017 level as shown in Table 11. Figures 5 and 6 are scatter plots that show the lack of correlation.

Table 11
Pearson Correlations (r) of LPCI scores correlated to Conceptual Understanding Determined by Interview

LPCI	Conceptual Understanding Pre-Instruction (n=24)	Conceptual Understanding Post-Instruction (n=24)	Conceptual Understanding Gain (n=24)
Pre-Instruction	0.613(0.001)		
Post-Instruction		0.436(0.033)	
Gain			0.078(0.718)

Figure 4

Interview Conceptual Understanding Prior to Instruction Plotted as a Function of LPCI Scores Prior to Instruction (n=24)

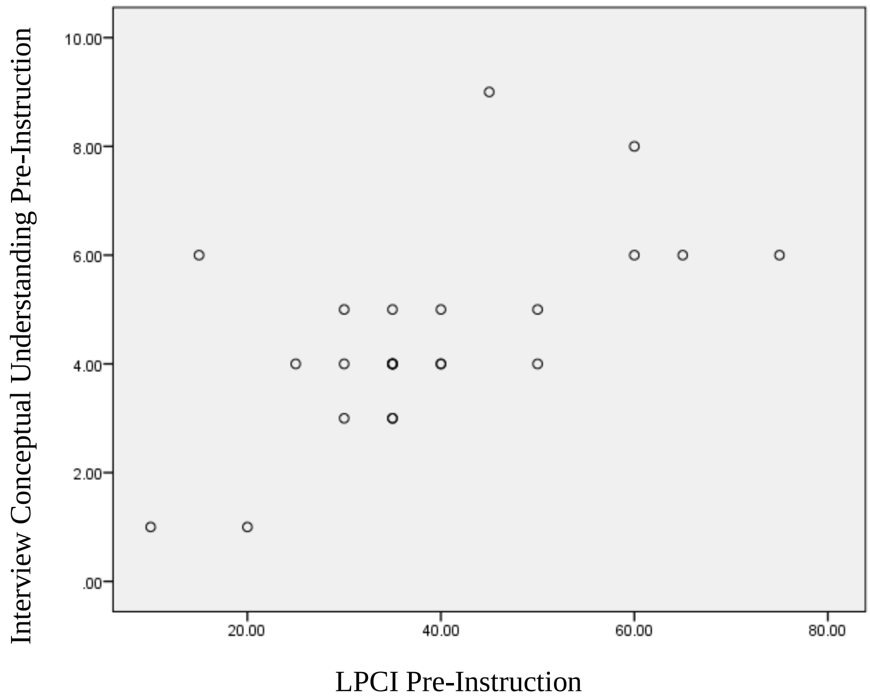


Figure 5

Interview Conceptual Understanding After Instruction Plotted as a Function of LPCI Scores After Instruction (n=24)

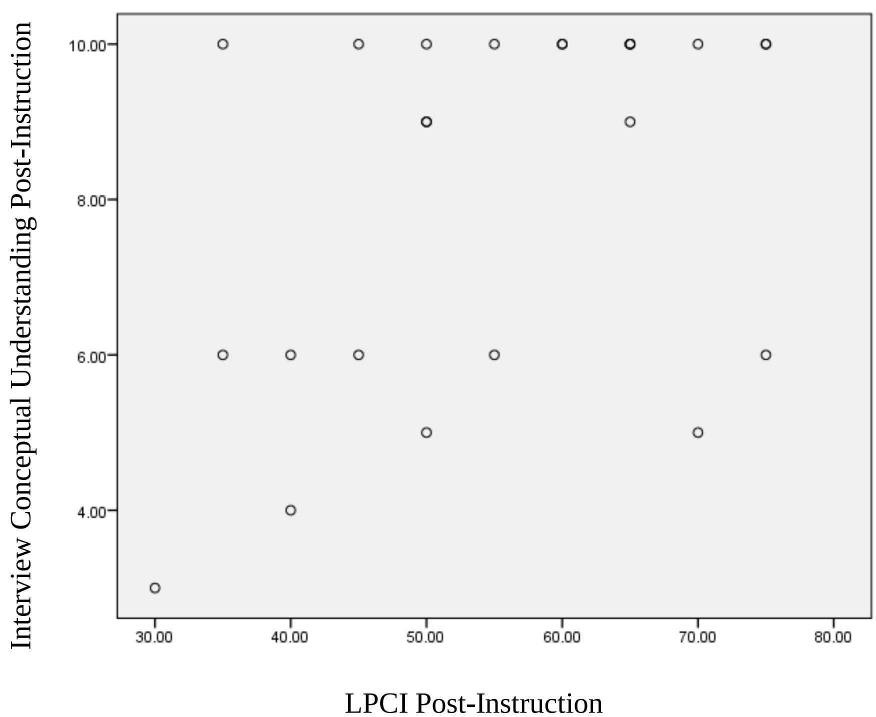
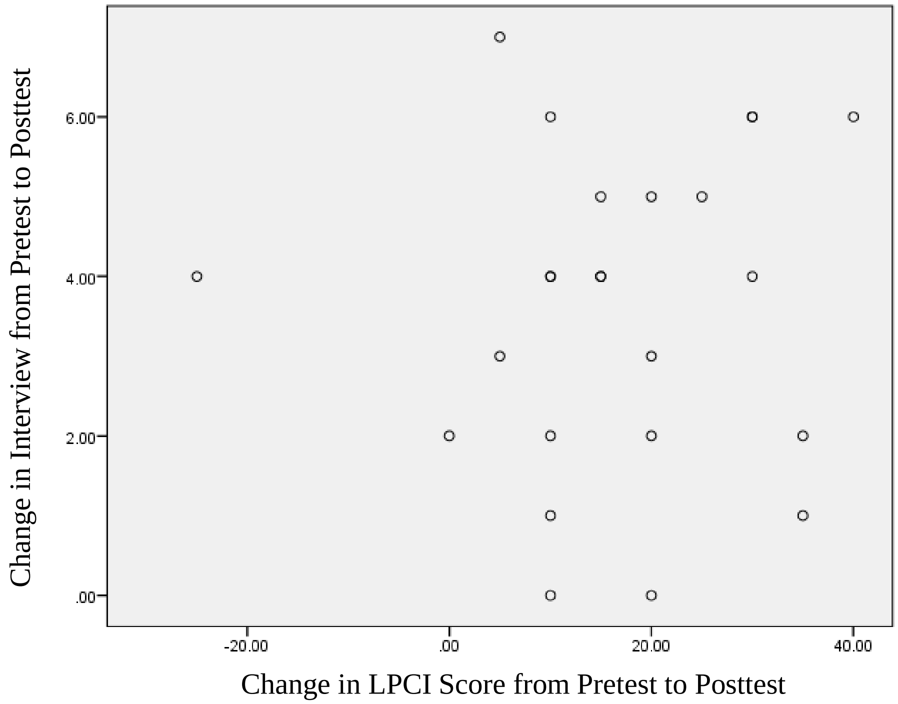


Figure 6
Change in Interview Conceptual Understanding Plotted as a Function of Change in LPCI Scores (n=24)



Correlations of Abridged-LPCI Scores to Interview Results

Table 12 shows that the Abridged-LPCI for the pre-instruction is correlated with the pre-instruction interview and is statistically significant ($r = 0.676$, $p \approx 0.000$) and that the Abridged-LPCI for the post-instruction is correlated with the post-instruction interview and is statistically significant ($r = 0.540$, $p = 0.006$) at the alpha = 0.017 level. As above, the modified-alpha level of 0.017 was used after the alpha of 0.05 was modified using the Bonferroni correction for three tests so that the adjusted alpha = 0.05/3 = 0.017. The statistically significantly correlation of the pre-instruction interview and pre-instruction Abridged-LPCI is shown by scatter plot in Figure 7 and the statistically significantly correlation of the post-instruction interview and post-instruction Abridged-LPCI is shown by scatter plot in Figure 8. The correlation for Gains (post-pre) are not statistically significant at the alpha = 0.017 level as shown in Table 12. Figure 9 is a scatter plot that shows the lack of correlation for the Gains.

Table 12
Pearson Correlations (p) of Abridged-LPCI scores correlated to Conceptual Understanding Determined by Interview

Abridged-LPCI	Conceptual Understanding Pre-Instruction (*n*=24)	Conceptual Understanding Post-Instruction (*n*=24)	Conceptual Understanding Gain (*n*=24)
Pre-Instruction	0.676(0.000)		
Post-Instruction		0.540(0.006)	
Gain			0.271(0.200)

Figure 7
Interview Conceptual Understanding Prior to Instruction Plotted as a Function of Abridged-LPCI Scores Prior to Instruction (n=24)

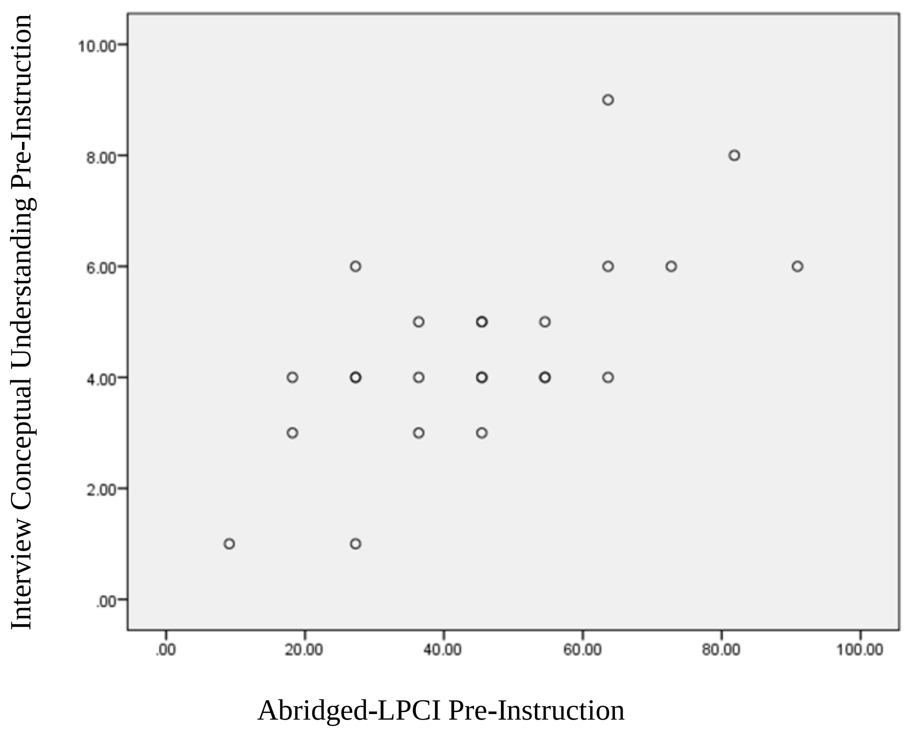


Figure 8

Interview Conceptual Understanding after Instruction Plotted as a Function of Abridged-LPCI Scores after Instruction (n=24)

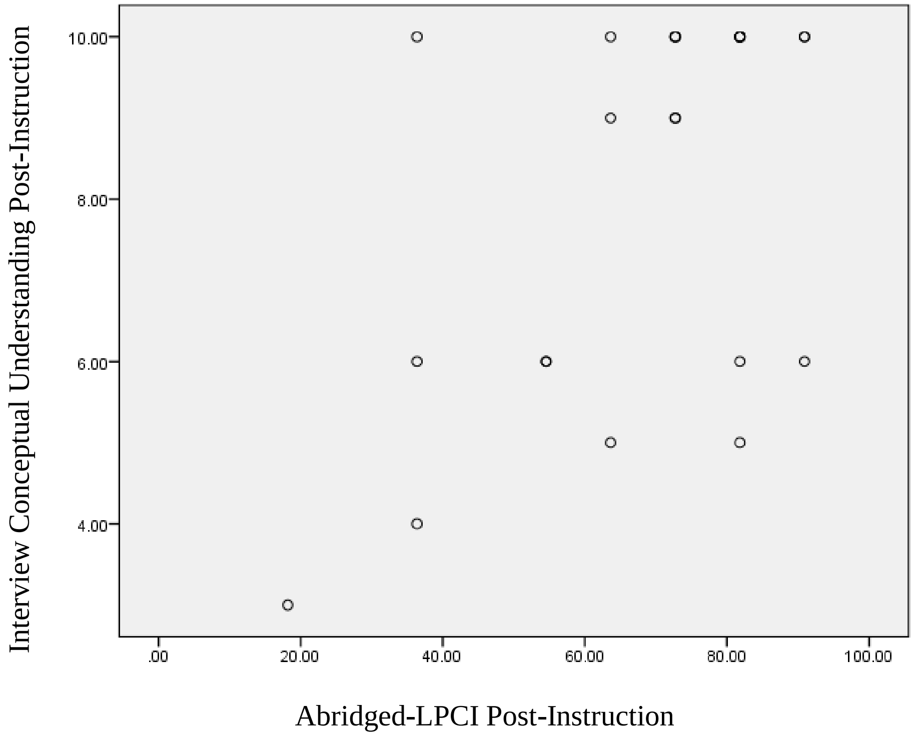
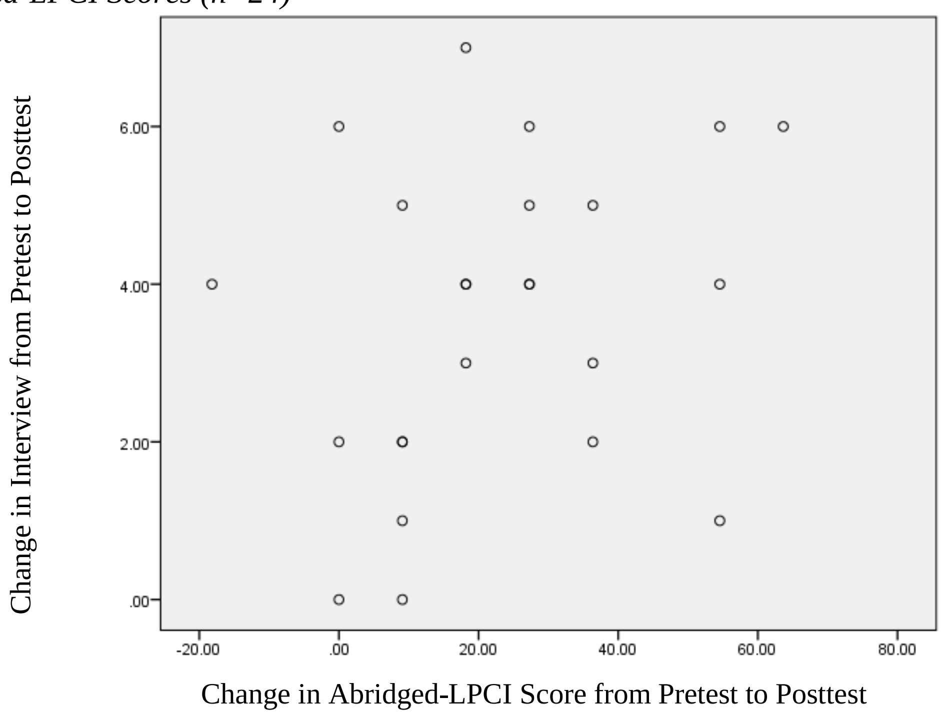


Figure 9

Change in Interview Conceptual Understanding Plotted as a Function of Change in Abridged-LPCI Scores (n=24)



Research Question #1 (Conceptual Change Effectiveness)

In answer to the first research question**:** How do pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases change after inquiry-based, conceptual change instructional intervention based on a computer simulation? It appears that there are substantial increases in conceptual understanding. For instance, according to Table 5, the change in conceptual understanding increased to 8.1(2.4) from 4.5(1.8), increasing 3.5(2.0) points on a 10 point scale (*n*=24) when measured by interview. Also from Table 5, the change in conceptual understanding when measured with the LPCI increased to 52.8(15.3) from 36.7(16.4) increasing 16.1(15.3) points on a 100 point scale (*n*=69). The change in conceptual understanding when measured with the Abridged-LPCI increased to 64.4(20.7) from 41.1(20.0) increasing 23.3(20.7) points on 100 point scale (*n*=69). The interview protocol resulted in a greater relative increase (3.5 points on a 10-point scale) than either LPCI (16.1 points on a 100-point scale) or the Abridged-LPCI (23.3 on a 100-point scale). Therefore, the Abridged-LPCI gains are in more agreement with the gains measured by the interview. However, the effect size for both forced-choice instruments and the interview protocol were all categorized as large. According to Table 5, all of the measures of scientific understanding increased after instruction for both semesters combined using interview data, LPCI, and the Abridged-LPCI. Additionally, all the measures of alternative conceptions decreased or stayed the same before and after treatment for both semesters combined using interview data in Table 6. Additionally, the treatment included participants taking data from new moon to the third quarter moon. This was effectively 21 days (evenings) of data and the remaining portion of the lunar cycle (third quarter back to new moon) was used for students to test their mental models. Therefore, 21 consecutive days (evenings) of data were sufficient to for students to build effective mental models.

According to Table 10, the time (pre-instruction vs. post-instruction) was significant for combined first and second semester LPCI data. According to Table 11, time was significant for combined first and second semester Abridged-LPCI data. Therefore, there are statistical and substantive differences (gains) in achievement when comparing post-instruction results to pre-instructions results for both LPCI and Abridged-LPCI measures. Therefore, pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases improve both statistically and substantively after inquiry-based conceptual change instructional intervention based on a computer simulation.

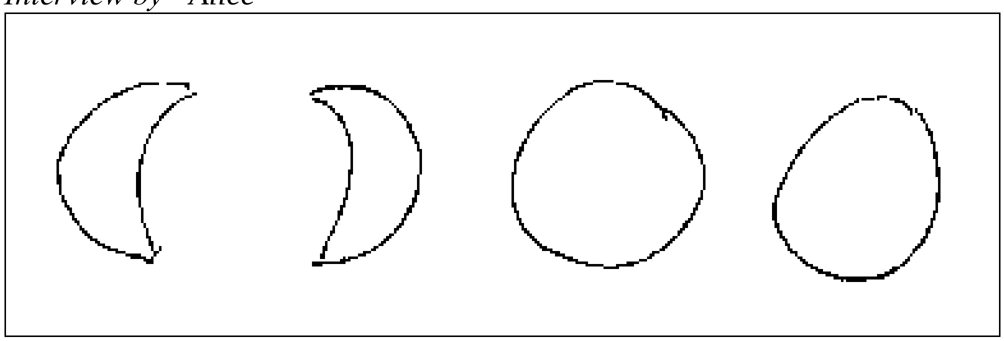
Conceptual Change from Interview Results

The interviews that were conducted with a purposeful sampling of each of the class sections can be useful to illuminate the conceptual change process as related to research question #1.   When the pre-treatment interviews are compared to the post-treatment interviews, changes in especially alternative conceptions can be documented.

Evidence for conceptual change can come from specific alternative conceptions being absent in post-treatment interviews. An example of such conceptual change came with the interview results for Alice (pseudonym). Alice's pre-instruction interview was scored at 4 points using the scoring rubric given in APPENDIX C and exhibited specific the alternative conception of moon orbiting the Earth within 24 hours. Additionally, Alice's LPCI score was 50% for pre-instruction. Alice's responses are somewhat contradictory since she mentions a monthly cycle and a daily cycle in the form of the 24 hour response. Additionally, Alice incorrectly sorted the lunar phase cards. Below is an excerpt from the pre-treatment interview with Alice.

> Interviewer: How would you explain the causes of the moon phases? What's going on that makes it [the moon] look different in the sky?
> *Alice: Um, here we go…well, it's based on like of course how it orbits around the Earth.*
> Interviewer: How what orbits?
> Alice: The moon, the moon is in between the Sun and the Earth.
> ***[Alice uses the scientific models to show the moon orbiting the earth.]***
> *Alice: I'm pretty sure the moon goes around the Earth, um… I can't, um… well I know that it goes around like, um, all the phases are within a month, but then I'm pretty sure that the moon orbits around the Earth in 24 hours, or the length of a day.*
> Interviewer: So you have the moon going around the Earth, in you said 24 hours?
> *Alice: I think so, yeah.*

Figure 10
Pre-Treatment Drawing Not Showing Earth-Moon-Sun Positional Data Made During Interview by "Alice"

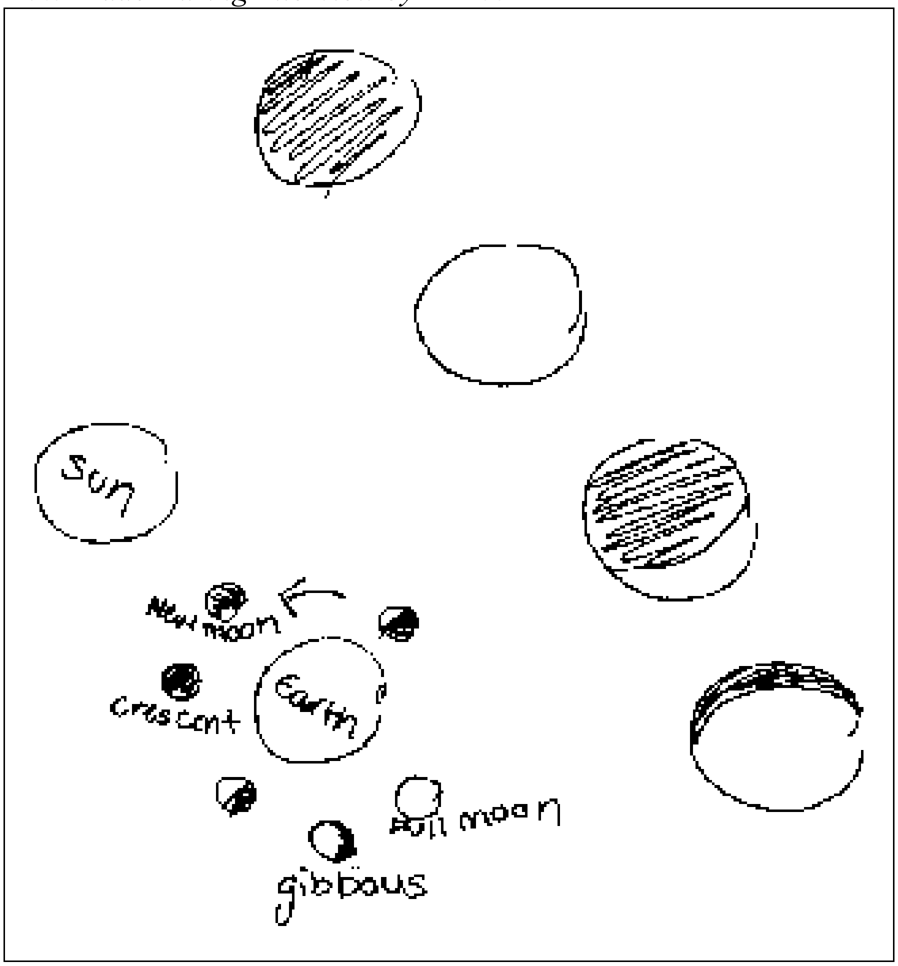


For the post-treatment interview with Alice, the 24 hour orbit of the Moon around Earth is missing from her explanation of the cause of the lunar phases. Below is an excerpt from the post-treatment interview with Alice.

> Interviewer: Um, so can you give an explanation of what causes the lunar phases?
> *Alice: The lunar phases are caused by, um, the Moon revolving around the Earth, because depending on the position between the Sun and the Earth, um, depending on the position of the Moon between the Sun and the Earth, depends on um, how much of the Moon we can see, like if the moon is between the Sun and the Earth, then it would be new moon, because the side that is lit up by the Sun is the opposite side of the moon and we can't see that. It's still there, we just can't see the shape of it necessarily. While if the moon is on the opposite side of the Earth, you can see, it will be a full moon, the Sun will be illuminating the side that we can see.*

[Alice correctly sorts the lunar phase cards in the correct order.]
Interviewer: Can you explain the cycle?
Alice: Ok, well, it starts off with the new moon at the beginning of the month. Um, I actually don't know if it's at the beginning of the month, but I know it takes a month, or twenty nine days to do a complete cycle.

Figure 11
Post-Treatment Drawing Showing Scientifically Correct Earth-Moon-Sun Positional Data Made During Interview by "Alice"



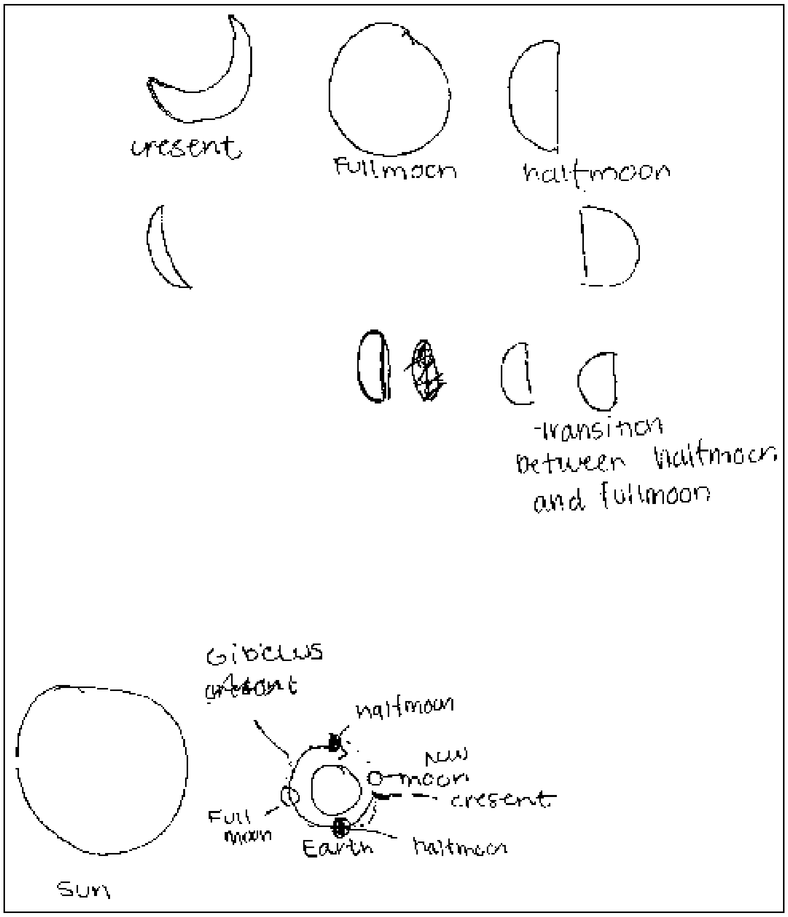
Alice correctly sorted the lunar phase cards and her interview was scored at 10 points for post-instruction. Additionally, Alice scored a 60% on the LPCI after instruction, compared to her initial score of 50% during pre-instruction. However, Alice's abridged-LPCI results were 64% for both pre and post instruction results, resulting in no change. Nevertheless, Alice's drawings improved from illustrating a limited number of lunar phases during the pre-instruction (Figure 10) interview to showing more scientific moon phases along with a representation of the relative positions of the Earth, Moon, and Sun necessary to scientifically explain the cause of the lunar phases during her post-instruction interview (Figure 11). The example of Alice illuminates the process of conceptual change in science because her alternative conception of the 24 hour cycle for lunar phases has been replaced by a monthly cycle to explain the lunar phases which she links in her post-treatment interview to the Moon's orbit around the Earth. This gives an example of a student constructing their own understanding using a constructivist approach by way of conceptual change with the use of *Starry Night* © software since data were taken over a 21 day period that resulted in observations of new moon to third quarter moon.

More evidence for conceptual change comes from the interviews conducted with Betty (pseudonym). Betty held the alternative conception that the Earth's shadow is the cause of lunar phases (AltEcl) during her pre-instruction interview, which was scored at 4 points using the rubric from APPENDIX C because she held only two of the four scientific concepts for lunar phases along with one alternative conception. She also incorrectly sorted the lunar phase shapes during the interview. Additionally her LPCI score was 40% for pre-instruction. Below is an excerpt from the pre-treatment interview with Betty along with the drawing she made during her interview; her drawing is shown in Figure 12.

> Interviewer: Ok, very good. Do these moon shapes occur in a predictable way, or a random way?
> *Betty: In a predictable way.*
> Interviewer: How long would it take for these shapes to repeat themselves in the sky?
> *Betty: Um, I would say about three weeks. The period of a month.*
> Interviewer: So between three weeks and a month?
> *Betty: Uh, huh.*

Interviewer: So if you had to explain to somebody what causes these shapes, what would you give as an explanation?
Betty: Um, I guess, the um, the shadow of the Earth… of the Sun, on the Moon.
Interviewer: So you've mentioned three things. We have the Earth, right, then we have the Moon of course, then we have the Sun. Ok, so let's say that… let's make a diagram how would the Earth and moon be situated with respect to the Sun?
Betty: Like the Sun here, then the Earth, and the Moon.
[Betty makes drawing of the relative positions of the Sun, Earth, and Moon]
Interviewer: So what do we know about the Moon and it's motion? What does the Moon do?
Betty: It goes around the Earth.
Interviewer: Ok, so maybe we could draw that.
[Betty draws circle around the Earth indicating the orbit of the Moon around the Earth and labels the Sun, the Earth, and the Moon]
Interviewer: So where would we put the moon for us to see a full moon from Earth?
Betty: Right here.
[Betty draws a full moon in between the Sun and Earth on the diagram]
Interviewer: How about the half moon and the crescent moon?
[Betty draws the half moon 90 degrees from the full moon, the crescent moon is drawn beyond the half moon]
Betty: Between here and here, would be a half moon, and here would be the crescent.
Interviewer: Ok, go ahead and label those. How about the transition between full moon and half moon? Let's go ahead and call that a gibbous moon.
[Betty draws the position of a gibbous moon between the half moon and the full moon]
Interviewer: So what would the moon right across from the full moon be? Is there another shape we haven't listed yet?
Betty: Um, like a "no moon".
Interviewer: A "no moon" or maybe a new moon? It's Just another name for it.
[Betty draws new moon opposite the full moon]
Interviewer: So we have some three dimensional models here. We have the Sun, the Moon, and the Earth.
Betty: Okay.
Interviewer: Could you position the moon so that it would be full?
[Betty positions the moon model between the models of the Sun and Earth]
Betty: So like right there.
Interviewer: How about the new moon, where would that be?
[Betty positions the moon opposite the full moon]
Interviewer: How about the crescent moons and the half moon?
[Betty positions the crescent and half moons according to the diagram]
Interviewer: And how about the gibbous moons, where would they be?
[Betty positions the gibbous moons according to the diagram]
Betty: Um, in like the transitional area, like right here.

Figure 12
Pre-Treatment Drawing Showing Earth Eclipsing the Moon to Produce Moon Phases and Over and Under Articulated Crescent Moon Made During Interview by "Betty"



Betty held the alternative conception that phases of the moon are caused by the Earth casting a shadow on the moon (AltEcl). This can be further verified by the over-articulated crescent moon shown by the top crescent moon drawn by Betty in Figure 12. At the bottom of Figure 12 Betty drew the new moon on the opposite side of the Earth than the Sun, which further illustrates the AltEcl alternative conception.

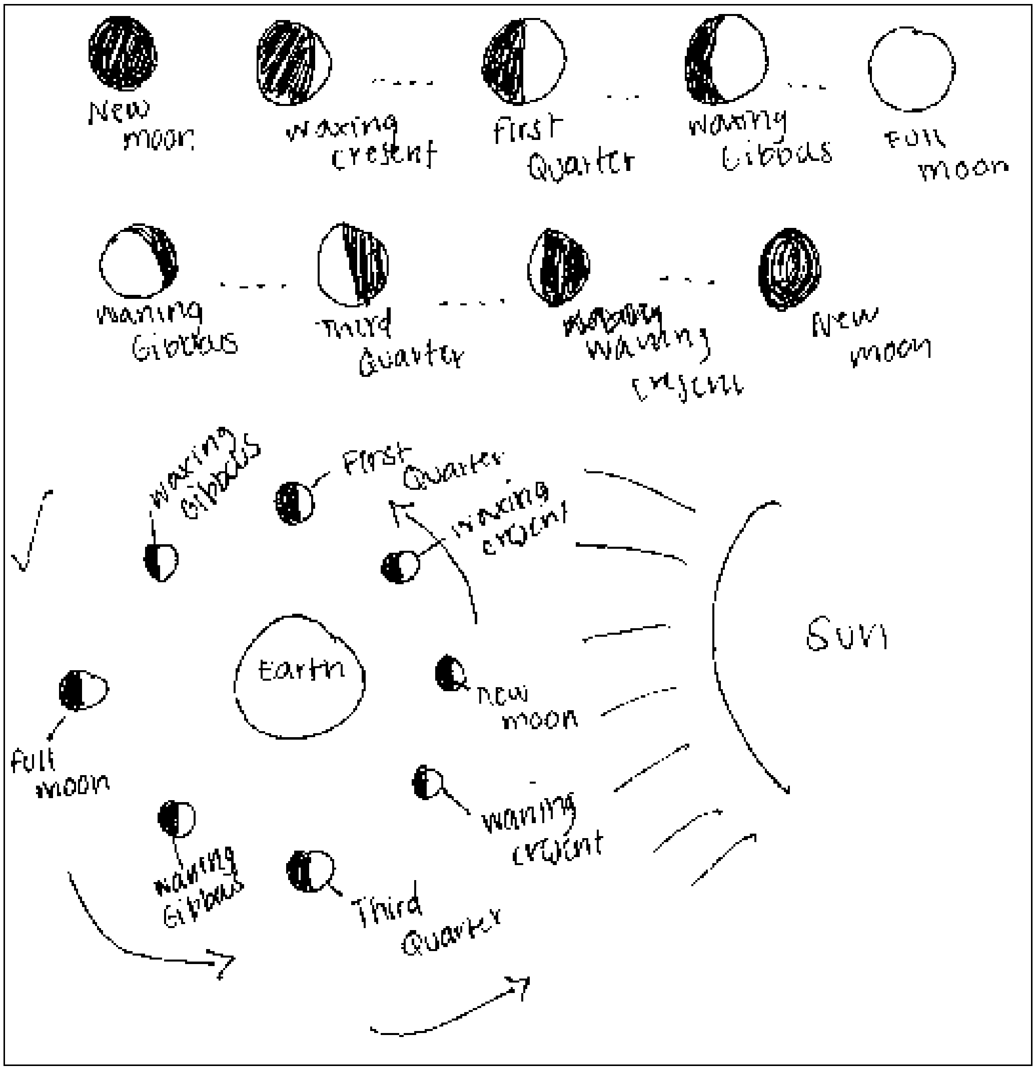
For the post-treatment interview with Betty, the alternative conception of the Earth casting a shadow onto the Moon to produce lunar phases (AltEcl) is missing from her explanation and is replaced with the scientific explanation. Below is an excerpt from the post-treatment interview with Betty along with the drawing she made during her interview along with the drawing she made during her post-instruction interview as shown in Figure 13.

> Interviewer: Thanks for coming in. Some of these questions will sound familiar to you. So, when I mention lunar phases, what comes to mind?
> *Betty: Like the stages of the moon throughout the course of a month.*
> Interviewer: And so how does the moon go through changes or stages?
> *Betty: As it revolves around the Earth.*
> Interviewer: Ok, so what happens to the Moon throughout the month?
> *Betty: Um, it starts off as dark, and then it gets a little bit bigger throughout the course of the month.*
> Interviewer: So, is it the shape of the moon that's changing?
> *Betty: Well, no, it's not the shape of the moon that's changing; it's how much that we can see.*
> Interviewer: Ok, could you draw some of these changes that we see?
> ***[Betty draws and labels the entire cycle from including all phases]***
> Interviewer: So, if you had to explain to somebody what causes these changes, how would you explain that?
> *Betty: Um, I'd say it's the Sun, and the relationship of the moon in between the Earth and the Sun.*
> Interviewer: Ok, could you maybe make a diagram?
> *Betty: Sure, here is the Earth, and put the Sun over here, and we'll put the new moon right here. And it starts to orbit this way. And the full moon is over here.*
> ***[Betty draws and labels the positional diagram that shows the entire cycle from including all phases]***
> Interviewer: How would you explain the shapes using this diagram?
> *Betty: I was really confused at first, because when we did the activity in the class with the Styrofoam™ ball on the Popsicle™ stick, we were turning and then when my body was like when the moon was in front of me, I was creating darkness onto it, then I realized that I'm standing up straight and the Earth is at a tilt. So that's why the Sun is able to reach over to the moon because of the Earth's tilt.*
> Interviewer: How would you draw in the sunlight on the moons?
> ***[Betty draws that only the side of the moon facing the sun is illuminated]***
> *Betty: Like this, I know they're not half moons, it's just how the sun is reflecting.*
> Interviewer: Ok, good. So how would you explain the first quarter moon?
> *Betty: Um, half the moon is being lit up by the sun and the other side is dark.*
> Interviewer: How would you explain the third quarter moon, how would you explain the difference in shape?
> *Betty: Well, I mean you're facing a different direction on the Earth.*
> Interviewer: I'm going to bring out these three dimensional models. This is the Sun, this is the Moon and this is the Earth. Can you place these in a relationship to each other that would explain the new moon?
> ***[Betty places the moon in between the Earth and Sun]***

Interviewer: Can you explain what's happening to the light from the sun.
Betty: The sun is basically just hitting the back side of the moon. So when we're here on Earth, we're not seeing anything lit up.
Interviewer: How about the full moon?
[Betty places the moon on the opposite side of the Earth compared to the Sun]
Betty: The full moon would be over here, um, and I mean, the Earth is at a tilt, so the Sun's you know obviously a lot larger, so it's being able to light up this whole entire side.
Interviewer: So how do you explain the full moon not being blocked by the Earth and the Moon being in the Earth's shadow?
Betty: Um, I think it's because of the Earth's tilt.

Betty's post-instruction interview was scored at 6 points because although she held all four scientific concepts she communicated one new alternative conception related to the tilt of the Earth, or AltETilt. Her LPCI score improved to 75% compared to her 40% LPCI score for pre-instruction. Additionally, Betty's abridged-LPCI results increased from 55% for pre-instruction to 91% for her post-instruction results. Additionally, her sorting of lunar phases was correct. The example of Betty illuminates the process of conceptual change in science because her alternative conception the Earth's shadow causing the lunar phases has been replaced by the four scientific concepts albeit with a new alternative conception understanding in her post-treatment interview. Never the less, her post-instruction interview was scored at 6 points using APPENDIX C which is an increase from 4 points for her pre-instruction interview. The changing of her conceptual understanding is graphically illustrated in her drawings of Figure 12 and 13, most notably the lack of the Earth casting a shadow for her post-instruction interview drawing of Figure 13. Additionally, Figure 13 does not have an overly articulated crescent as was present in Figure 12 that indicated the AltEcl alternative conception was held. However, Figure 13 still exhibits an under-articulated crescent shapes which may be indicative of the AltETilt alternative conception.

Figure 13
Betty's Post-Treatment Drawings Showing Scientifically Accurate Model of Moon Phases, Scientifically Correct Gibbous Phases, and No Overly-Articulated Crescent Moon



Conceptual Change from Student Reflection Results

In addition to the interview results, the student reflections can also illuminate the conceptual change process as related to research question #1. As part of the treatment, the students confront their alternative conceptions and replace them on the last day of instruction. This takes the form of students evaluating their conceptual models and revising them with the use of *Starry Night* © software or the use of their 3-D scientific models.

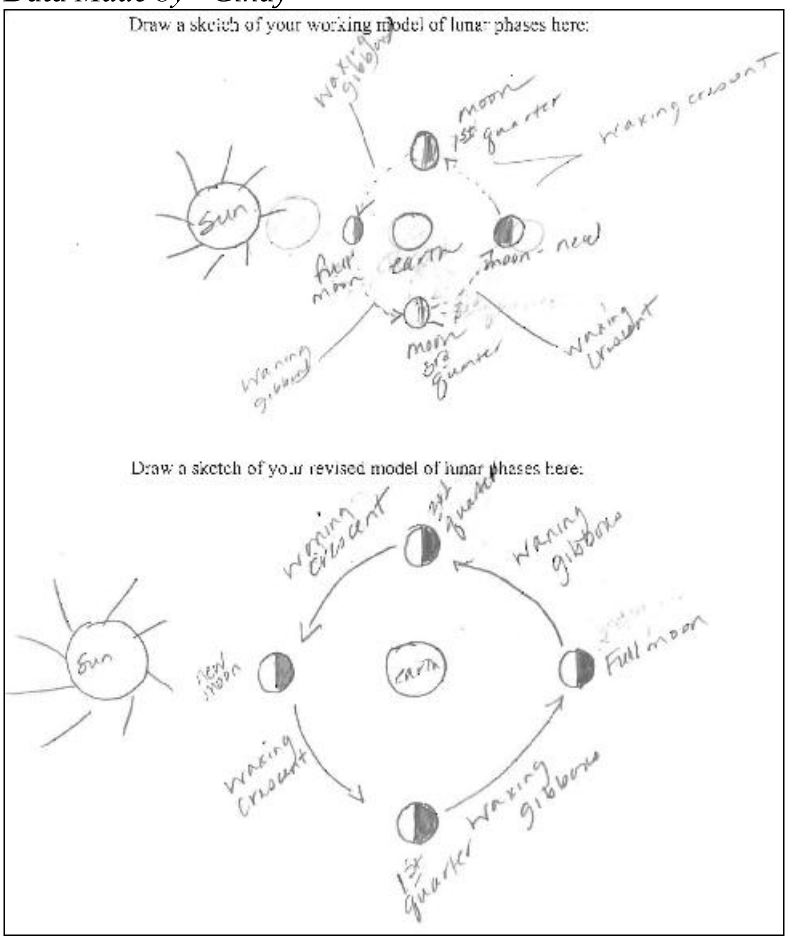
The student reflection kept by Cindy (pseudonym) gives a good example of a student revising their conceptual model because they have confronted their alternative conception with data they have collected. Cindy held the alternative conception that the new moon was caused by the Earth's shadow which is illustrated with initial conception where she placed the new moon on the opposite side of the Earth from the Sun. Cindy correctly revised her model which used the scientific explanation of the lunar phases. Cindy states that the use of the 3-D scientific model (Styrofoam™ ball activity) was most helpful for her. The reflection was scored with a high level of engagement using the rubric from Chapter 3. Cindy's pre-instruction LPCI score was 70% and her post-instruction LPCI score was 60%. Although her score went down by ten points, her 60% score still indicates a working knowledge of lunar phases and compares well with the average of 52.9% as shown in Table 5. The 70% result from pre-instruction was much higher than the average of 36.7% which is also shown in Table 5. Therefore the reduction from 70% to 60% when pre-instruction is compared to post-instruction could be due to the regression toward to mean and may not be a true indication of student conceptual gains. This regression toward the mean was not found for Cindy's abridged-LPCI results however. Cindy's abridged-LPCI results were 73% for pre-instruction and 82% for post-instruction. The data from the reflection demonstrates that her understanding improved after instruction, although the LPCI results may indicate the contrary. Below is an excerpt from Cindy's reflection along with the drawing she made (Figure 14) of her original and revised mental models.

> I was thinking backwards related to the illumination of the moon for new and full moons. I thought that the new moon would be 180 degrees from the sun and the full moon would be closest in its orbit to the sun. However, it is opposite- the sun is 180 degrees from the moon when the moon is at full moon-we can see the side

of the moon that will be facing the sun. At new moon we are looking at the side of the moon facing the earth but turned from the sun. At this point, the moon is closest to its orbit to the sun. I was able to predict closely the last few days of the moo's orbit and illuminated by looking at its waning pattern from full to 3rd quarter and its waning crescent phases. Using the Styrofoam ball and the light was most helpful as a visual to understand the way the sun illuminates the moon in its different phases.

Figure 14

Student Reflection Drawing Showing Conceptual Change of Earth-Moon-Sun Positional Data Made by "Cindy"



In Figure 14 Cindy demonstrates that she has replaced the AltEcl alternative conception because the new moon and full moon positions are switched when comparing initial mental model with the revised mental model.

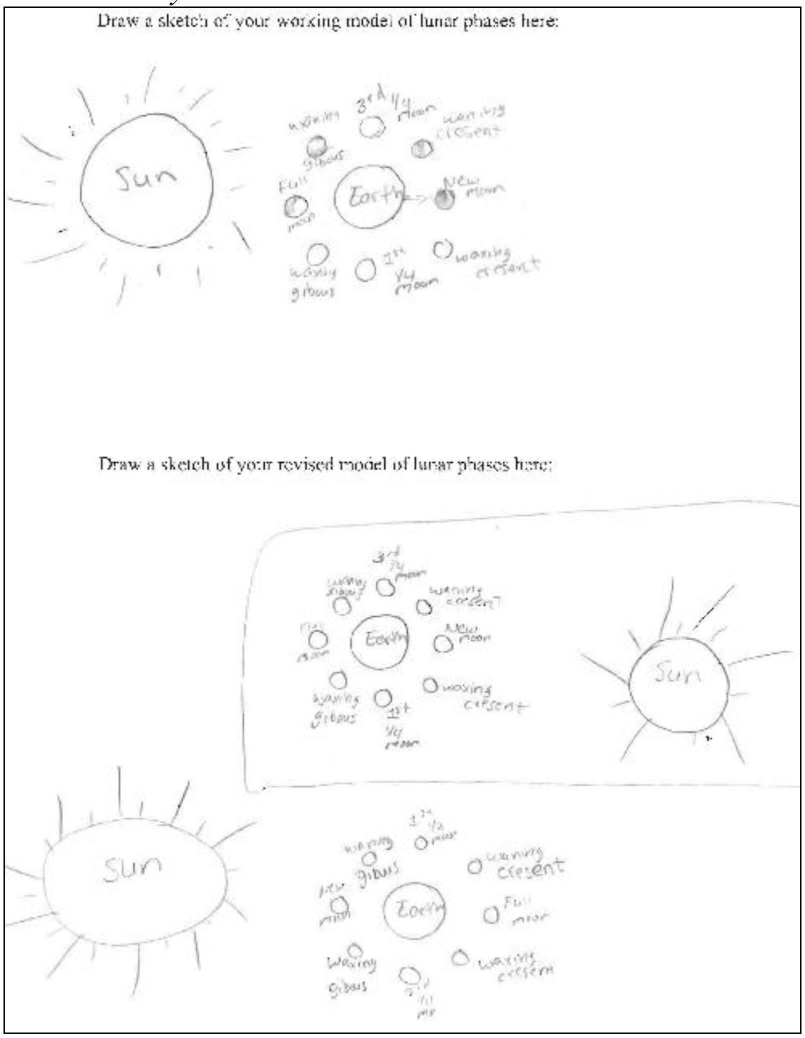
Diane's (pseudonym) reflection provides another example of a student revising their conceptual model because they have confronted their alternative conception with data they have collected. Diane, like Cindy, held the alternative conception that the new moon was caused by the Earth's shadow which is illustrated with initial conception where she placed the new moon on the opposite side of the Earth from the Sun. Diane also correctly revised her model which used the scientific explanation of the lunar phases. The reflection was also scored with a high level of engagement using the rubric from Chapter 3. Diane's pre-instruction LPCI was 45% and her post-instruction LPCI resulted in a score of 65%, thus indicating an improvement in conceptual understanding of lunar phases. Similarly, Diane's abridge-LPCI results for pre-instruction were 64% and her post-instruction abridged-LPCI result was 73%. In this case the LPCI results agree with the conceptual change indicated by her reflection data. Below is an excerpt from Diane's reflection along with the drawing she made (Figure 15) of her original and revised mental models.

> The last day of this project really made everything all come together. When we went over where the sun, moon, and earth are in realitive [*sic*] to each other. My thoughts about waxing and waning did not change. However I was reminded that a half the moon, waxing, is a first quarter moon and a third quarter when it is waning. I think this helped me in my understanding of the positioning of the moon, sun, and earth.

Results from both the student interviews, as well as daily student reflections, can give insights into conceptual change in science education that is not available with the forced-choice instruments such as the LPCI and the Abridged-LPCI. Namely, how do the participants replace their alternative conceptions with scientifically accurate concepts when then their original working models are confronted with data. These interviews and reflections show that conceptual change can be used to improvements in concepts of shape (Figure 11), sequence (Table 7), and causes of moon phases (Figures 13, 14, and 15) among pre-service elementary teachers.

Figure 15

Student Reflection Drawing Showing Conceptual Change of Earth-Moon-Sun Positional Data Made by "Diane"



Research Question #2 (comparison of treatments)

In answer to the second research question: Are there differences in achievement between pre-service elementary teachers who collect moon observation data in a whole-class setting versus those who collect this data working in pairs with the computer simulation?, there appears to be no statistical or substantial differences. For instance, according to Table 9, the treatment was non-significant for the LPCI data. According to Table 10, the treatment was non-significant for the Abridged-LPCI data. Therefore, there are no statistical or substantive differences in achievement between students who collect moon observation data in a whole-class setting versus those who collect this data working with the computer simulation in pairs when using LPCI and Abridged-LPCI data.

Research Question #3 (contribution of instructional components)

Correlations of Reflections to LPCI Scores

Table 13 shows that word count is correlated with LPCI gain and is statistically significant ($r = 0.338$, $p = 0.004$) at the alpha = 0.017 level. The modified-alpha level of 0.017 was used after the alpha of 0.05 was modified using the Bonferroni correction for 3 tests so that the adjusted alpha = 0.05/3 = 0.017. The three tests come from the three correlations of LPCI results with word count measurements. Additionally, the level of engagement is correlated with the LPCI post instruction results and is statistically significant ($r = 0.314$, $p = 0.009$) at the alpha = 0.017 level. The other measures of the student reflection (number of questions and number of predictions) are not statistically significantly correlated with any of the LPCI measures (Pre-Instruction, Post-Instruction, and Post-Pre).

Table 13
Pearson Correlations (p) of LPCI scores correlated to Student Reflection measures

LPCI	Word Count (*n*=69)	Number of Questions (*n*=69)	Number of Predictions (*n*=69)	Level of Engagement (*n*=69)
Pre-Instruction	-0.077(0.532)	0.099(0.417)	-0.110(0.366)	0.188(0.122)
Post-Instruction	0.257(0.033)	0.240(0.047)	-0.087(0.480)	0.314(0.009)*
Post-Pre	0.338(0.004)*	0.132(0.279)	0.032(0.793)	0.112(0.360)

* = Indicates statistically significant Pearson Correlation

Correlations of Reflections to Abridged-LPCI Scores

Table 14 shows that none of the measures of the student reflection (word count, number of questions, number of predictions, and level of engagement) are statistically significantly correlated with any of the Abridged-LPCI measures (Pre-Instruction, Post-Instruction, and Post-Pre) at the alpha = 0.017 level.

Table 14
Pearson Correlations (p) of Abridged-LPCI scores correlated to Student Reflection measures

Abridged-LPCI	Word Count (*n*=69)	Number of Questions (*n*=69)	Number of Predictions (*n*=69)	Level of Engagement (*n*=69)
Pre-Instruction	-0.030(0.806)	0.168(0.168)	-0.011(0.930)	0.179(0.140)
Post-Instruction	0.246(0.041)	0.213(0.079)	-0.034(0.779)	0.229(0.059)
Post-Pre	0.275(0.022)	0.050(0.681)	-0.024(0.846)	0.055(0.653)

Student-Perceived Usefulness of Instructional Components

Data from Table 15 show that 70.8% of the participants responded that the Styrofoam™ ball (scientific model) activity was the most important instructional element. This factor was much greater than the other three responses combined. The *Starry Night©* software elicited only four responses, making drawings only had one response and the response that making drawings and the scientific model activity being equally important had two responses. Therefore, the scientific model activity was by far the most significant instructional factor as reported by the participants.

Table 15

Most Important Instructional Factor as Reported by Students during Interview

Factor	*n*	Number	%
Scientific model Activity	24	17	70.8
Starry Night© Software	24	4	16.7
Making Drawings	24	1	4.2
Making Drawings and Scientific model Activity Equally Important	24	2	8.3
Would Use Same Type of Instruction with their Future Students	24	23	95.8

Table 16 is a summary of the students that would use the same type of instruction in their own classrooms, how would they change the instruction so as to customize it for their own classrooms. The data reflect a follow up question asked during the post-instruction interview. Almost half (47.8%) reported that they would make no changes to the instruction. Less than one fifth (17.4%) of the students responded that they would not use *Starry Night*© Software. A small number of students (13.0%) responded that they would prefer to have the “last day” activities sooner in the instructional sequence. This means that the students prefer to develop mental models, revise mental models, and make use of the scientific model activity sooner in the instructional sequence. Although a majority of the students that would use this type of instruction in their own future classrooms, there is not a single conclusive change that the students would recommend.

Table 16
How Students Would Change Instruction in their Future Classrooms

Change	*n*	Number	%
Make no changes	23	11	47.8
Would not use *Starry Night*© Software	23	4	17.4
"Last Day" of instruction sooner	23	3	13.0
Shorten overall number of lessons	23	2	8.7
Make real sky observations	23	2	8.7
Make more drawings	23	1	4.3

Reflections

The data in Table 17 show that 26.1% of students offered the unsolicited response that the scientific model activity was useful. Moreover, 8.7% (6 out of 69) or the participants gave unsolicited responses that the *Starry Night*© software was confusing, compared to only one participant offered that that software was helpful. This is in contrast to post-instruction interview data from Table 15 that indicates a large majority of students, 95.8% (23 out of 24), reported that they would use this type of instruction with their future students. Of the students that would use this type of instruction in their own future classrooms, only 17.4% (4 out of 23) stated that they would not use *Starry Night*© software. Therefore, the students did not have a significant or consistent adverse view of the use of *Starry Night*© software. Additionally, a majority of participants (71%) drew under-articulated crescent moons in their reflection, 2.9% drew overly-articulated crescent moons and 30.4% drew false gibbous moons. These scientifically inaccurate drawings are an indication that the students may hold alternative conceptions regarding lunar phases.

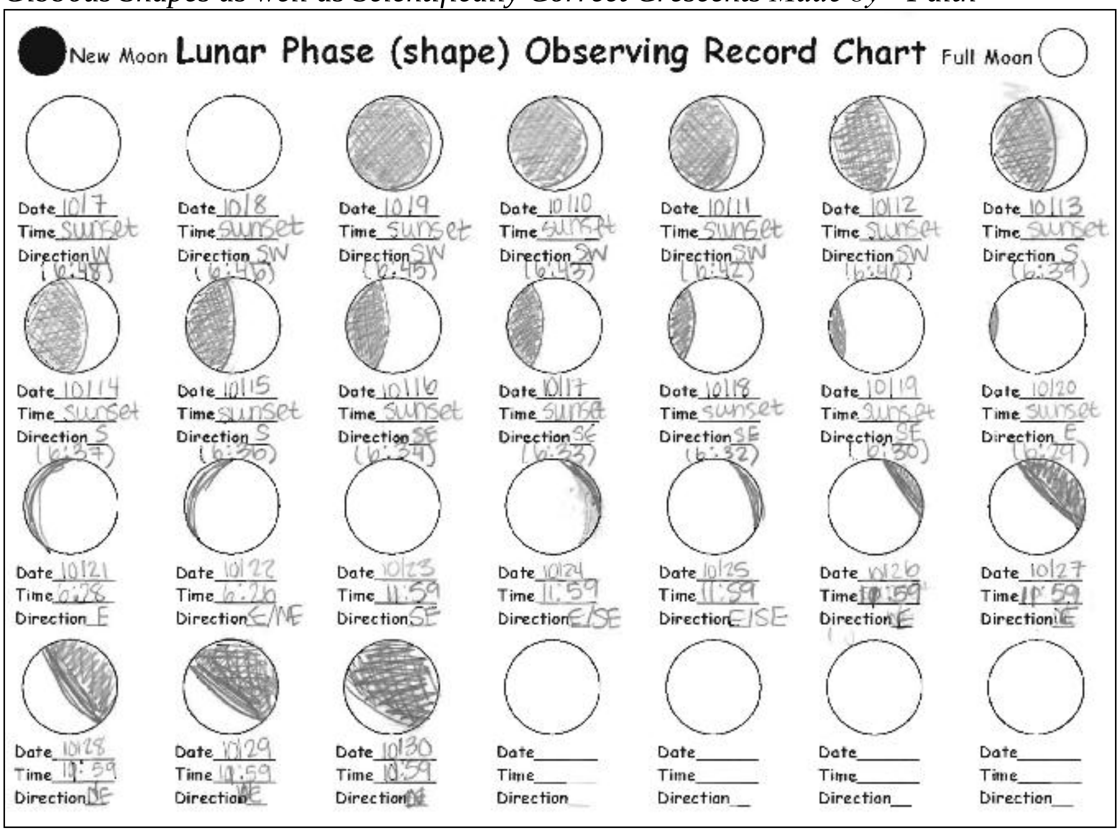
Table 17
Frequencies of Student Responses from Student Reflection

Response	*n*	Number	%
Scientific model was useful	69	18	26.1
Software was confusing	69	6	8.7
Software was helpful	69	1	1.4
Drew under-articulated crescent	69	49	71.0
Drew overly-articulated crescent	69	2	2.9
Drew false gibbous	69	21	30.4

Below (Figure 16) is given an example of student drawings made in their student reflection. In this example Faith (pseudonym) draws the shapes of the lunar phases of observations using *Starry Night©* that she took during instruction. There are several examples of alternative drawings of lunar phases given. Figure 16 shows the drawing for 10/10 is an example of an overly-articulated crescent. Likewise, an example of an under-articulated crescent is given for the dates 10/29 and 10/30. Additionally, examples of the "false gibbous" are given for dates 10/15-10/20. A good example of a waxing crescent is illustrated for 10/11. In the top two rows of Figure 13, Betty illustrates and labels scientifically accurate examples of waxing and waning gibbous phases of the moon.

From Table 17, there are a relatively large number of participants that drew under-articulated crescents (71.0%), fewer drew overly-articulated crescents (2.9%), and a moderate number (30.4%) drew a false gibbous in their reflections. The significant number of participants making alternative drawings of moon phases does not correspond with only 3 out of 24 students responding that making drawings was an important factor for instruction. Although 70.8% of the students reported that the scientific model activity was the most important activity during treatment, the use of making drawings may be more important than reported by the participants as evidenced by the large number of drawings that illustrate alternative conceptions of lunar shapes.

Figure 16
Student Reflection Drawing Showing Overly and Under Articulated Crescents and False Gibbous Shapes as well as Scientifically Correct Crescents Made by "Faith"



Below are responses from student reflections that are examples of the most important instructional factors are reported by students.

> [Gina (pseudonym)] The program was confusing but when we did the activity with the styrafoam [*sic*] balls and the light bulb I got it.

> [Helen (pseudonym)] I feel like I better understand the phases of the moon after this final activity and being able to interact with objects to better understand the phases of the moon.

> [Isabel (pseudonym)] After finally finishing the moon observation phases lesson, I feel confident that I understand the change of the moon and phases as it rotates [revolves around] the Earth in perspective of the sun. With the ball on the popsicle [*sic*] stick and the light bulb, I found it really understandable to actually see how the light works in view of the sun and the moons position to me, the earth.
>
> [Joni (pseudonym)] The Styrofoam ball activity helped me to really understand the shape of the moon and what it relies on, the sun. I think that spreading the data collection over many days made understanding the phases much harder. If we did the data collection all at once, that would have accelerated the understanding.

The quotes from the student reflections raise some interesting points related to conceptual change. Isabel states that she "found it really understandable to actually see how the light works" and Helen states that "to interact with object to better understand the phases of the moon". Both of these quotes illuminate the process of conceptual change by students being able to manipulate a model (in this case a physical model) in order to revise and replace an initial mental model with a scientific mental model. Students Gina and Joni also state that the scientific model activity was very helpful for their understanding, and therefore their conceptual change. Joni further recommends that she would have preferred gathering data over fewer days.

The data from Table 18 show that 29.2% (7 of 24) participants in the post-instruction interview responded that the reflection was not important while 70.8% (17 or 24) responded that the reflection was "a little" important. Although 70.8% of the students reported that the scientific model activity was the most important activity during treatment, the use of making drawings during the writing of the daily reflection may be more important than reported by the participants as evidenced by the large number of drawings that illustrate alternative conceptions of lunar shapes as shown in Table 16.

Table 18
Role of Writing a Reflection during Instruction as Reported by Students during Interview

Student Response	*n*	Number	%
Reflection was not important	24	7	29.2
Reflection was "a little" important	24	17	70.8

To answer the third research question: How do pre-service elementary teachers perceive the relative value of each component of the instructional intervention as related to their attainment of scientific conceptions of moon phases? According to Table 13, the level of engagement measured by the student reflections, has a statistically significant correlation with post-instruction LPCI scores. Also according to Table 13, the word count as measured by the student reflections has a statistically significant correlation with LPCI gain (post-pre) scores. Although there were statistically significant correlations, they were not practically significant. However, according to Table 14, none of the factors measured by the student reflections had a statistically significant correlation with any factors of the Abridged-LPCI. In general, the factors measured by the student reflections are not strongly correlated with the LPCI (2 out of 12 possible correlations) and none of the Abridged-LPCI factors are correlated with any of the student reflection factors (zero of 12 possible correlations).

According to Table 15, over two-thirds (70.8%) of the interview participants stated that the scientific model activities were the most important instructional factor. Moreover according to Table 16 the scientific model activity was also mentioned in 26.1% of the reflections as being useful. This was by far the most prevalent useful component mentioned. According to Table 18, zero participants found student reflections were significantly helpful, 29.2% stating that the reflection was not important, and 70.8% stating that the reflection was “a little” important.

Therefore, the scientific model activity was perceived by the participants as the most significant factor that positively impacted improvement of scientific conceptions of moon phases. The use of *Starry Night* © software was also perceived by students as being another important instructional component of the treatment. There is evidence that writing a daily reflection which includes drawings may be important than participants report as evidenced by the prevalence of scientifically inaccurate drawings illustrated in student daily reflections. The word count of the student reflections is correlated with LPCI gain and is statistically significant. Additionally, the level of engagement is correlated with the LPCI post instruction results and is statistically significant. There were no statistically significant correlations measured for Abridged-LPCI results. However, although the correlations are statistically significant, they are not practically significant and give no strong indications that student reflections are correlated with LPCI or Abridged-LPCI results.

Research Question #4 (Comparison of LPCI to semi-structured interview)

In answer to research question number four: How does the forced-choice LPCI instrument compare with the semi-structured Trundle protocol in measuring the conceptual understanding of lunar phases among pre-service elementary teachers? The concept domains of the LPCI and the elements of the semi-structured interview of the Trundle protocol are not directly aligned as shown in Figure 1. Nine of the twenty questions of the LPCI were removed in order to develop the Abridged-LPCI in order to better represent the topics covered in the semi-structured interview protocol. The pre-instruction results of the LPCI were correlated with the semi-structured interview when administered before instruction and the correlation was statistically significant ($r = 0.613$, $p = 0.001$) at the alpha = 0.017 level. The Abridged-LPCI showed statistically significant correlations between the pre-instruction ($r = 0.676$, ≈ 0.000) and post-instruction ($r = 0.540$, $p = 0.006$), both at the alpha = 0.017 level, with the corresponding semi-structured interview results. Although statistically significant, Pearson correlations of 0.613, 0.676, and 0.540 result in a coefficients of determination, or the "R-squared" values, of 0.376, 0.457, and 0.291 respectively. These coefficients of determination give the portion of the fluctuation of the LPCI (or Abridged-LPCI) that can be predicted by the interview, or vice versa. So, although statistically significant, they describe the majority of the variability and thus, do not substitute well for each other.

Additionally, both the LPCI and the Abridged-LPCI are not able to measure important insights into the conceptual change process as evidenced by the interviews with participants "Alice" and "Betty". Therefore, to answer research question number four, the Abridged-LPCI is a better choice than the LPCI to replace the semi-structured interview when only a measurement of conceptual understanding is needed. However the use of the forced-choice instruments can be problematic because they cannot measure critical elements of the conceptual change approach to science teaching.

CHAPTER 5: DISCUSSION AND IMPLICATIONS

Introduction

The present study was successful in gaining insights into how pre-service teachers learn moon phase concepts, which results in recommendations for the classroom as well as future research topics. These recommendations include continued use of an inquiry-based, conceptual change approach to instruction for lunar phases that incorporates the use of planetarium software to gather data. Future use and research into the scientific model activity is also encouraged for classroom practice as well as for future research. The use of scientific models to teach the Earth-Moon-Sun system to K-5 students is specifically recommended in the *Next Generation Science Standards*, (National Research Council, 2013). These recommendations mention not only describing the motions and relative positions of the Earth, Moon, and Sun, but also using the scientific models to make predictions. The present study made use of the scientific models to not only describe the lunar phases but to also make predictions of lunar phases into the future.

For the present study, pre-service elementary teachers' conceptions of shape, sequence, and cause of moon phases improve both statistically and substantively after inquiry-based, conceptual change instructional intervention based on a computer simulation. These students were able to test and revise their mental models, often based on alternative conceptions, using data they collected. The students then replaced their previous mental models with scientifically accurate mental models. Additionally, there is a consistent result that differences due to the treatment, namely comparing demo of *Starry Night*© versus participants working in small groups, are not statistically significant. Therefore, there are no statistical or substantive differences in achievement between students who collect moon observation data in a whole class setting versus those who collect this data working with the computer simulation in pairs. The use of daily student reflections was found to be useful to gain insights into student learning, such as students finding that the scientific model activity being useful, but is not necessarily recommended because there are not connections between student reflections and conceptual gains. Data gathered from the interviews and student reflections served to illuminate the conceptual change process when students received instruction. And finally, the scientific model activity was found to be a significant factor contributing to student learning of lunar phases and further research is warranted.

Findings

Pre-service elementary teachers' conceptions of shape, sequence, and scientific understanding of the cause of moon phases improved both statistically and substantively after inquiry-based, conceptual change instructional intervention based on a computer simulation. Gains in conceptual understanding were measured by a semi-structured interview protocol as well as the LPCI and an Abridged-LPCI that is a version of the LPCI that includes only the questions that align directly with the interview protocol. The Abridged-LPCI results correlated better than the LPCI with the semi-structured interview protocol, which was not unexpected since the abridged-LPCI specifically only used questions from the LPCI that were aligned with the concepts addressed by the semi-structured interview protocol. The key finding is that the abridged-LPCI represents a reasonable alternative for the semi-structured interview since the interview protocol is much more labor intensive than administering the forced-choice abridged-LPCI instrument, although statistically significant, the correlations are not substantially significant. Additionally, the use of the forced-choice instruments can be problematic because they cannot measure critical elements of conceptual change.

Also, there were no statistical or substantive differences in achievement between students who collected moon observation data in a whole class setting versus those who collect this data working with the computer simulation in pairs. In general, the factors measured from the writing of student reflections were not significantly correlated with measures of conceptual gains although they can provide insights into the learning process.

And finally, the scientific model activity was found to be the most significant instructional component that positively impacted improvement of scientific conceptions of moon phases. Only 2 of 12 possible correlations between student reflection and LPCI were statistically significantly correlated and none of the 12 possible correlations for the Abridged-LPCI were statistically significantly correlated with student reflections. Although there were statistically significant correlations, they were not practically significant. Moreover, the students generally reported that reflections were not an important factor in their attainment of scientific conceptions of moon phases. Additionally, three new alternative lunar phase alternative conceptions were identified in pre-instruction interviews. These new alternative conceptions include: Alternative Orbital Speed (the moon orbits the Earth in 24 hours, thus causing the lunar phases), Alternative No Orbit (the moon is in the same position relative to the Earth and Sun for all of the lunar phases), and Alternative Large Sun (the Sun is larger than the Earth, so light goes around the Earth causing the full moon). It was found that the semi-structured interview protocol was more able to measure a variety of alternative conceptions relating to lunar phases concepts when compared to a forced-choice instrument such as the LPCI or the Abridged-LPCI.

The above findings have implications for teaching practice using computer simulations and conceptual change for both practicing and pre-service teachers. Additionally, many new lines of research can result from this present research which can include: longitudinal studies, the use of scientific models, as well as the exploration of demographic data.

Discussion

Inquiry-Based, Conceptual Change Instruction

The present study showed that inquiry-based, conceptual change instruction for lunar phase concepts can result substantial gains in conceptual understanding. For instance, the number of participants holding a fully scientific understanding of lunar phases was found to increase from 17% (4 of 24) to 71% (17 of 24) after instruction using the semi-structured interview protocol. Additionally, the LPCI and the Abridged-LPCI results also showed significant gains. This finding is supported by Parker and Heywood (1998) where it was found that repetitive exposure to a scientific explanation (didactic teaching) does not guarantee a scientific understanding of astronomical events. The treatment in the present study was based on McDermott (1996) was similar to Trundle et al. (2002, 2007a, 2007b, 2010) and resulted in similar gains in conceptual understanding of lunar phases.

A common element in all of treatments from these studies is the necessity of providing guided inquiry where the instructor facilitates the learning process by providing prompts when necessary but allowing students to develop and evaluate their own mental models. Scaffolding was also used to provide learners with the appropriate conceptual context in order for them to develop their own mental models. This inquiry-based, conceptual change instruction using scaffolding was specifically designed to promote conceptual change where learners have to confront their alternative conceptions and replace them with scientific concepts. This approach was also found to be very important in Alparslan et al. (2003) where learners confront their alternative conceptions and are dissatisfied and then a replacement is made with the scientific conception. The Alparslan et al. (2003) study was particularly applicable since it concentrated on theoretical model construction and evaluation. More importantly Windschitl and Andre (1998) also found that computer simulations can offer an appropriate environment for college students to test their mental models during constructivist instruction for science topics.

Revising and testing of mental models was also an important consideration in Keating et al. (2002) when investigating the learning of college students. However, other researchers (Barnett & Morran, 2002, p. 875) showed that "instruction does not necessarily need to directly address students' alternative framework to promote conceptual change". However, Barnett and Morran (2002) further discuss that their study consisted of only 17 self-selected fifth grade participants so their results are not generalizable. Therefore, Barnett and Morran (2002) should not be considered because their study cannot be generalizable to the present study. The scaffolding in the present study includes both formal and informal classroom discussion which supported the cognitive dissonance required for students to confront their alternative conceptions. This type of classroom discussion was found to be important in Trumper (2000) when investigating university students. Kavanagh et al. (2005) additionally recommends the use of this "constructivist" approach to learning theory for lunar phases, as opposed to direct instruction which is unable to engage learners' old mental models and construct new mental models.

The present study used a treatment that had the participants gather 21 days (three weeks) of data from new moon to third quarter moon. Most of the studies related to the present research had participants gathering nine weeks of data. These studies included Trundle et al. (2002, 2006, 2007a, 2007b, 2010), Bell and Trundle (2006, 2007), Trundle and Bell (2010). However, Bell et al. (2007) and Bell and Trundle (2008) had participants gather data for 16 days and four weeks respectively. All of these studies showed similar significant conceptual gains for lunar phase concepts.

Specifically, Bell and Trundle (2008) found that 82% of participants held a scientific view of lunar phases after instruction using a conceptual change treatment that that used four weeks for data collection and used two additional class periods to analyze the data. Similarly, the present study found that 71% of participants held a scientific understanding of lunar phases using a similar conceptual change approach, however using 21 total days. Thus the present study supports gathering data for a shorter time period, in this case 21 days, can result in similar significant gains in scientific understanding of the lunar phases.

The use of the conceptual change approach to science education was found to be useful for the present study. The key element and theoretical framework of the conceptual change approach was when participants are confronted with disconfirming evidence related to their alternate conception(s), ideally they are forced to replace this (these) alternative conception(s) with a scientifically accurate conception(s). This revision, or wholesale replacement, of mental models during instruction was also a key finding in Vosniadou and Brewer (1994). This process is difficult to document with the use of forced-choice instruments because they are only administered before and after instruction and rely on only a “right” or “wrong” scoring for each question which are limited by the number of distractors used. In order to better understand the learning process of conceptual change, data can be taken during the learning process by means of daily reflections taken by the learners. These reflections should include open responses in addition to drawings to illustrate their working mental models. Additionally, interviews should be conducted with participants before and after instruction that can lead to more open responses related to lunar phases. These more open responses facilitate conceptual understanding of lunar phases to be more fully expressed by the participants than is possible with a forced-choice instrument because the participants’ responses are not limited by pre-established distractors.

The interviews and student reflections help to illuminate the process of conceptual change that the students experience as part of the treatments. The interview results demonstrate that students can replace their alternative concepts with scientific concepts to explain the reason for lunar phases, although not every student achieved the same level of conceptual understanding. The use of student drawings can help to illustrate these conceptual changes. Students look for patterns and cycles in order to develop their mental models. Student drawings from the daily student reflections also indicate a similar replacement of alternative conceptions with scientific concepts. Therefore, the present study demonstrates that data taken from interviews and reflections should be used to investigate how learners replace their alternative conceptions when confronted with disconfirming scientific observations during the conceptual change process.

Use of Planetarium Software

The findings of this study have implications for both the teaching of pre-service teachers, but also to the methodologies of future studies. Results indicate that inquiry-based use of planetarium software is an effective approach to teaching lunar phases to pre-service teachers through use of conceptual change. The use of planetarium software in conjunction with an inquiry-based, conceptual change approach to teaching was found result in substantial and significant gains in scientific understanding of lunar phases. Interview results indicate that this conceptual change approach allowed participants to construct their own knowledge with the use of *Starry Night* © software. This finding is consistent with other studies such as: Bell and Trundle (2006, 2007, 2008), Bell et al. (2007), Trundle and Bell (2010), and Hobson et al. (2010). As an example, Bell and Trundle (2008) showed that 82% (pretest = 0%) of participants held a scientific understanding of the cause of the phases of the moon after instruction, whereas the present study showed that 71% (pretest = 17%) of the participants held a fully scientific view after the instructional treatment. In a similar way, Trundle and Bell (2010) showed that 80% (pretest = 4%) of participants held a fully scientific view after instruction when the treatment was planetarium software only. These findings are echoed in Monaghan and Clement (1999), Rutten et al. (2012) and Smetana and Bell (2012). Although the present study does not result in 100% of students holding a scientific understanding of lunar phases, in fact, some participants revert back to their alternative-conceptions after only one day after instruction. A finding that participants revert back to their alternative conceptions was also reported in Trundle et al. (2007b).

The above findings strengthen the case that planetarium software, such as *Starry Night©* should be used for teaching lunar phases concepts as part of an inquiry-based, conceptual change approach, especially for pre-service teachers since the present results are consistent with past studies. The present study therefore indicates that it is not necessary for learners to make observations of the natural moon in order to obtain significant gains in conceptual understanding, as is consistent with past studies. Similarly, Trundle and Bell (2010) find no differences between the use of planetarium software only for lunar observations when compared to natural observations only as well as compared to a mixture of natural sky observations and data taken with the use of planetarium software. Additionally, Smetana and Bell (2012) in a review article find that computer simulations can be as effective, or even more effective, than traditional instruction methods.

The use of planetarium software in conjunction with an inquiry-based conceptual change approach to teaching the lunar phases should also be considered by pre-service teachers for their own teaching. This is due to the effectiveness of this approach either with a whole-class or students working in pairs approach. Additionally, a large majority of students, 95.8% (23 out of 24), reported that they would use this type of instruction with their future students during their post-instructions interviews as presented in Table

15. As shown in Figure 16, of the students that would use this type of instruction in their own classrooms, only 17.4% (4 out of 23) responded during their interview that they would not use *Starry Night©* software but would use the other instructional activities. Additionally, these pre-service teachers should consider the availability of software when planning their own lessons. Only software that is generally available is recommended for use, so that educators can plan to make use of their lesson plans for future years.

Whole Group versus Small Group

The findings of the present study show that there are no statistical differences between small group (students working in pairs at one computer) and whole group (instructor using one computer for demonstration) instruction. This finding was echoed by Smetana (2008) in a study that investigated the use of computer simulations to facilitate the learning of chemistry concepts which compared whole-class and small-group settings. Although the Smetana (2008) study investigated high school students, the results can be generally compared with the current study because the study also investigated student use of computer simulations for science teaching. Additionally, Chang (2003) found that teacher-directed computer assisted instruction resulted in greater gains than the student-directed computer assisted computer assisted instruction. Chang (2003) also investigated high school students.   Similarly, Hogan et al. (2000) found that teacher-guided instruction led to higher levels of reasoning for middle-school students. Conversely, Tessier (2007) found that peer instruction in small groups led to improved student learning compared to traditional lectures among pre-service elementary school teachers enrolled in a biology class. Tien et al. (2002) also finds that peer-led, small group instruction is superior to whole-class instruction when investigating undergraduate college students. However, the Tessier (2007) and Tien et al. (2002) studies were not researching computer simulations which may require more scaffolding than simple peer instruction can provide. Additionally, there were pedagogical differences in these two studies compared to the present study. For instance, the treatment for Tessier (2007) included a modified traditional lecture approach. Similarly, Tien et al. (2002) compared traditional lectures to a workshop approach to teaching organic chemistry. Therefore, Tien et al. (2002) and Tessier (2007) cannot be directly compared to the present study. In the present study, in light of the current and previous research, both whole-class and small-group settings are appropriate for the use of computer simulations for lunar phases. Additionally, for classrooms with limited resources the use of planetarium software as a demonstration should be encouraged unless a specific learning environment would include specific technology goals that would be met if students are operating the software themselves to gather data.

Scientific model activity

Data collected in the present study point to the scientific model activity as the most important component of the instructional intervention. This finding is based not only on the instructional component being the most prevalent instructional component mentioned in the post-instructional interviews, but also on it being mentioned the most often in reflections written by all the participants. Hansen et al. (2004) suggests that for topics such as space concepts where spatial understanding is required, 3-D modeling appears to be advantageous over static or 2-D models because it fosters spatial reasoning among college students. Although both the present study as well as Hansen et al. (2004) both use computer simulations, the Hansen et al. (2004) study appears to use computer tools that simulate 3-D models more effectively. Monaghan and Clement (1999) also found that computer simulations can facilitate mental simulations. Additionally, Parker and Heywood (1998) found that college students have difficulty visualizing a 3-D phenomenon, such as lunar eclipses, when a 2-D representation is used. The *Starry Night©* program in the present study showed motion of the moon in the sky along the plane of the sky and therefore the computer simulation was displaying effectively a 2-D (plane of the sky) image, not a 3-D simulation. In this study, the use of the scientific model allowed the students to vary their perspective and line-of-sight and thus perform a 3-D simulation of the Earth-Moon-Sun system. These findings were consistent with the literature.

The changing of perspective was also found to be critically important in Keating et al. (2002) in order to facilitate the conceptual change model of learning. Additionally, Vosniadou (1994) found that conceptual change can be particularly difficult to achieve when alternative conceptions exist with pre-suppositions of a conceptual framework. The present study suggests that this is the case for concepts related to lunar phases, since there are still a significant number of participants holding alternative conceptions of lunar phases after instruction (8 of 24 holding the alternative eclipse alternative conception, four of which while also correctly describing all four scientific criteria during their post-instruction interviews, and 1 of 24 holding the Earth's alternative tilt alternative conception, which also correctly describing all four scientific criteria during their post-instruction interview). Therefore, the ability of participants to develop their mental models in 3-D was critical. Moreover, the participants were able to instantly vary their viewing angles in order to quickly adjust their mental models as new data in the form of varying illumination of the scientific model became available. Kavanagh et al. (2005) recommends in a review article the use of a ball and lamp to simulate the moon and sun in order to facilitate learning of lunar phase concepts. Kavanagh et al. (2005) goes on to specifically not recommend the use of a flashlight because not all learners could see the same phase at from the same different locations because each learner's angle was different.

The present study suggests that each learner being able to manipulate their own model is an important factor for conceptual change. Excerpts from student daily reflections can illuminate the process of conceptual change by students being able to manipulate a model (in this case a physical model) in order to revise and replace an initial mental model with a scientific mental model when confronted with disconfirming data related to their alternative conception. Results from the student reflections also indicate that it is important for students to be able to see the models from differing positions. These findings indicate that the use of the scientific model activity should be encouraged when teaching pre-service teachers and should also be considered by pre-service teachers for their own teaching. The scientific model activity can be used in conjunction with the use of planetarium software or observations of the natural sky. Additionally, the scientific model activity is inexpensive and can be easily incorporated into classroom practice.

Use of Forced-Choice Instruments

The Abridged-LPCI was found to correlate both before and after treatment, whereas the LPCI only correlated before treatment with the interview. The Abridged-LPCI correlating both before and after treatment is likely due to eliminating questions from the LPCI not addressed in the interview. This result is not unexpected. Although there were statistically significant correlations, they were not practically significant, since the "R-squared" values are low. These differences are not unexpected because the interview protocol of Trundle was developed for use with pre-service elementary teachers and the LPCI was developed for undergraduates taking a traditional introductory astronomy class. The LPCI assesses a deeper understanding of lunar phase concepts, whereas the interview protocol investigates minimum understanding of lunar phase concepts that future elementary school teachers will need for their own teaching. However, if one were to use a forced-choice instrument, the Abridged-LPCI and the LPCI have an advantage over the use of a semi-structured interview in that these forced-choice instruments take much less time to administer. Nehm and Schonfeld (2008) found similar results with college students when the Conceptual Inventory of Natural Selection (CINS) was compared to an open response as well as an interview instrument. Namely, the CINS was found to be an excellent replacement of the interview when measuring conceptual understanding. The Nehm and Schonfeld (2008) recommendation was made in order to provide a quicker method to measure conceptual understanding of natural selection.

The Abridged-LPCI and the LPCI, in additional to lacking practically significant correlations compared to the interview protocol, also have a disadvantage that it is not possible to obtain a rich understanding of the participants conceptual understanding of lunar phases, and possibly more importantly, they are limited in their ability of measuring a fully variety of alternative conceptions that may be held by students. This was also the case for Nehm and Schonfeld (2008) where they found that the CINS produced significantly different results compared to the interview instrument when measuring the diversity and frequency of alternative conceptions. However, since practitioners seldom have time to conduct full interview with their students, and since there is some correlation with the Abridged-LPCI and the interview protocol, the Abridged LPCI can be used in place of the interview protocol in order to include larger sample sizes in future studies when only a measurement of conceptual understanding is needed, especially if some sub-set is interviewed so as to check the efficacy of the Abridged-LPCI for a specific situation. However evidence from the present study shows that the use of the forced-choice instruments can be problematic because they cannot measure critical elements of the conceptual change approach to science teaching.

Use of Reflections

The word count of student reflections are positively correlated with gains in conceptual understanding as measured by the LPCI and the level of engagement of the student reflections is positively correlated with the post-instruction LPCI results. However, there were no statistically significant correlations with the student reflections and any measure of the Abridged-LPCI. This was also found in Ruiz-Primo and Li (2004) where there were no correlations between student notebooks with learning outcomes, largely because the entries were found not to be "coherent". Although the Ruiz-Primo and Li (2004) study investigated notebooks of high-school age students, their findings can be compared to the present study because many of the reflections of the present study were also exhibited a low engagement. Conversely, Hyers (2001) found that academic measures were found to correlate with journal grades for college-aged learners enrolled in an introductory Earth Science course. In the Hyers (2001) study the journals were scored by the frequency of writing, word count, and the relevancy of the entry, although the journals were not scored according to the scientific accuracy of the entries. The literature is mixed regarding whether or not the use of notebooks and reflections by students is positively correlated with student outcomes. Although the present study found some correlations between student reflections with student outcomes, these correlations are not practically significant. However, there are no indications that the use of student reflections hinder student learning. Indeed, there may be benefits beyond gains in conceptual understanding of lunar phases, namely the instructor can gain insights into the teaching and learning process directly from the student reflections.

Abell et al. (2002) also stresses the importance that students continually revisit their moon learning experience by keeping a moon notebook. However, Abell et al. (2002) lacks specific criteria on how student notebooks were evaluated. The recommendation of Abell (2002) for pre-service teachers to keep a notebook during moon phase instruction lacks empirical evidence and should be carefully considered. Although, Parker and Heywood (1998) made use of annotated drawings and written reflections for their study of pre-service teachers and their understanding of basic astronomical events, they made no specific recommendations for use of these instruments. Additionally, Parker and Heywood (1998) state that pre-service teachers should be made aware of their own learning in order to improve their own teaching in the future. So, there may be reasons for pre-service teachers to keep reflections during instruction of lunar phases that are outside the scope of the measurement of student conceptual understanding. However, reflections are not generally recommended when time management is a consideration in the classroom and the reflections would possibly displace other content in the course. For instance, Cropp (1980) found that journals provide a critical venue where college students can provide feedback to the instructor about the course.

In the present study, many participants found that the scientific model activity helpful. Additionally, insights were gained regarding the conceptual change process. These were discovered by analyzing student reflections. However, the pre-service elementary teachers perceive that scientific model activity was more helpful than the writing of daily reflections. In fact, many more participants (18 out of 69) found the scientific model activity helpful as compared to no (0 out of 69) participants responding that it was helpful to write a daily reflection. This result was very similar to the findings from the interview where no (0 out of 24) participants stating that writing the reflection was at least moderately important. So moreover, student reflections, journals, notebooks and the like can give an instructor specific insights to the instructor about the course, independent of any possible research questions.

Applicability to Other Areas of Science Education

The findings of this study could also be applicable to other areas of science education that require students to visualize and/or make 3-D observations over time. Possible related areas could be: meteorology, chemistry, physics, and anatomy education research. For instance, in meteorology education, it may be beneficial to simulate the formation of various types of clouds over time as seen from the ground, or perhaps some other perspective. In chemistry education, perhaps teaching approaches that have students performing the simulation of the bonding, symmetry, and vibrational modes of molecules over time could benefit from findings of the present study. In physics education, possibly teaching approaches that have students learning projectile motion could benefit from these results. For anatomy, if a teaching approach includes the visualization of blood flow through the body, it might benefit from the present results. In general, any area of science education that benefits from students making observations over time, especially if those observations are difficult, or perhaps even impossible, the related teaching pedagogy could benefit from the replacement of alternative conceptions as facilitated by the *Starry Night*© program and/or scientific models as described in the present study. The computer simulation allows for the visualization of phenomena that can be difficult, if not impossible to observe in nature. Additionally computer simulations can allow for the manipulation of time as a variable. Time scales can often be changed easily, from minutes, to hours, days, years, to centuries, or even many millennia. Moreover, computer simulations can allow the student to vary time not only into the future, but also into the past.

Limitations

As described in Chapter 3, this study was conducted to mitigate threats to internal and external validity. The extent of internal validity will allow a case for causation to be made. The quality of the external validity will determine how applicable the findings of the present study are to other situations. The construct validity of both the LPCI and the interview protocol were addressed in the development of these instruments by other researchers.

Regarding internal validity, great care was taken during the study to insure that all treatments were equal except for the independent variables, which were the treatment used as well as time. Both instruments measuring the conceptual understanding of lunar phases showed similar results in conceptual gains. Moreover, the gains were consistent for both semesters. These results indicate that both instruments, the LPCI and the interview protocol along with the associated evaluation and coding of the interviews, were both measuring the students' conceptual understanding of lunar phases. There was 100% participation from each class and all the participants completed the study, so there were no biases towards only the high-performing students completing the study. These results suggest that the instruments were internally valid and causation can be attributed to the independent variables of treatment and time.

When considering external validity, the narrow demographics of this study limit how generalizable the results are to other groups of students. It was important to control for variables which the researcher had some control over. To this end, care was taken during the study to insure that the assignment of individuals into groups was as random as possible. All groups had similar pre-treatment measurements of conceptual understanding of lunar phases using the LPCI. The initial LPCI results were used to rank each class into a low, middle, and high scores. For each class, two students were randomly selected from each level for participation with the interview protocol. Each section of the class had equal number of instruction days. In order to insure external validity, these findings should be limited to undergraduate pre-service teachers enrolled a cohort introductory astronomy class in a similar undergraduate institution. These findings would be most applicable to small enrollment classes of 15-20 learners at a medium sized university with a strong program of science education specifically designed for pre-service teachers. Additionally, the present study was made with traditional college-aged students that were 97% female as well as 97% Caucasian.

The results from the present study can be compared to similar studies that investigated pre-service teachers (Trundle & Bell 2010, Bell & Trundle 2008). However, similar findings are also reported in studies investigating conceptual understanding of lunar phases of elementary school students (Trundle et al., 2007a, Hobson et al., 2010) as well as. A case can be made that findings reported in the present study could be extended to younger learners since similar results are obtained with both young learners, Trundle et al. (2010) and adult learners, Trundle et al. (2006). However, special care should be taken when making such extensions. For younger learners, a semi-structured interview protocol, such as is described in Trundle et al. (2006, 2007 a, b) and Hobson et al. (2010), would be a better than a forced-choice instrument. The use of the forced choice instruments such as the LPCI and the Abridged-LPCI are not generally recommended for younger learners. Moreover, Vosniadou et al. (2004) indicates that a forced-choice instrument limits the range of responses available to young learners and thereby limiting the formation of synthetic models.

Future Research

Instruments for Future Studies

The use of the Abridged-LPCI allows for larger scale studies into methods for teaching lunar phases. But just for studies that only require measurement of conceptual understanding since critical elements of conceptual change cannot be measured by the forced-choice instruments. This would allow assessments of classes with larger enrollments, even possibly including traditional introductory astronomy courses with small (10-30 enrollment), medium (30-70 enrollment), and large (70+ enrollment). One advantage to increasing sample size is that statistical analysis would be more robust and would allow for a wider variety of investigations. These varied investigations could address gender, race, socio-economic, learner preparedness differences as well as investigating differences between instructors.

The present study included only two male participants out of the total of 69 participants. Bisard et al. (1994) also found similar small numbers of males when studying pre-service teachers. The use of the Abridged-LPCI could therefore broaden the demographics of previous studies. The Abridged-LPCI can be alongside the Astronomy Diagnostic Test (ADT) as described in Hufnagel et al. (2000) and Hufnagel (2002), Zeilik and Morris (2003), and Deming (2002) which was also developed for use with college students. The LPCI as well as the Abridged-LPCI are measuring similar space science topics as the ADT. A combined Abridged-LPCI and ADT could provide useful assessment data for college-level astronomy courses for both pre-service teachers as well as a more typical undergraduate introductory astronomy course.

The ADT would be complimentary to the Abridged-LPCI because the ADT also gathers demographic data from the participants. The Abridged-LPCI would be better suited to combine with, or administer in conjunction with, the ADT because there are 9 less questions (11 instead of 20) on the Abridged-LPCI compared to the LPCI. The results can be either combined as one comprehensive assessment or given as a “part-A” and “part-B”. The demographic information from the ADT could be split off as a “part-C” and used with either parts (A or B) or (A and B).

Scientific Model Activity for Future Studies

An expanded study into the use of scientific models for lunar phase instruction should be undertaken. An issue that arises with the use of the scientific model activity is a learner's conceptual understanding of eclipses. Eclipses as an alternative conception of the cause of lunar phases among elementary school students was studied in Barnett and Moran (2002). Additionally, Kavanagh et al. (2005) indicated in a review article which investigated alternative conceptions at all level of students that confusion between eclipses and lunar phases are common among students. This issue has not been investigated in the literature specifically for the measurement of conceptual understanding of eclipses separate from the conceptual understanding of lunar phases. An investigation should study a learner's accommodation of eclipse concepts along with lunar phase concepts. This would include spatial manipulation of scientific models to show differences in the orbital planes of the moon compared to the orbital plane of the Earth around the Sun. The use of the scientific model activity shows promise as being an ideal modeling device for this study because it is easily manipulated in 3-D.

In the present study, the scientific model activity was performed by participants only on the last day of instruction that allowed the learners to summarize and synthesize their observation and make models and predictions. It has not been established how many days this activity should last. Participants in the present study gathered data 21 lunar days which was performed over a three-week period, lending support to a much shorter data collection period than the nine weeks specified by McDermott (1996). Perhaps gathering twenty one days of lunar data using *Starry Night©* during a single class session followed by the testing of mental models with 3-D models with the scientific model activity would be just as effective as longer treatments. A systematic investigation should be conducted that would measure the conceptual understanding of lunar phase concepts when the instruction occurs in one day and then compare those results to previous studies that used longer instruction times. If the shorter period of 1 day of instruction is equivalent to, or perhaps better than, longer periods, there would be the advantage of using only one day of instruction for lunar phase concepts. This would have the benefit of allowing more class time to be spent on other critical concepts related to space sciences.

Longitudinal Future Studies

Very few longitudinal studies have been conducted to investigate how tenaciously learners retain their conceptual understandings of lunar phase concepts. Trundle et al. (2007b) showed that the majority pre-service teachers continued their scientific understanding of lunar phases six months after instruction. Of course, it would be more interesting to investigate how tenacious pre-service teachers retain their scientific understanding of the cause of lunar phases after they become teachers and to see if they revert to their initial alternative conceptions related to lunar phases. It would also be of interest to investigate such teachers make use of inquiry-based, conceptual change instruction in their own classrooms that also makes use of planetarium software. An investigation that would study the conceptual understanding of lunar phase concepts of teachers several years after instruction would be a meaningful. With a sample size of 69 that took the LPCI and a subset of 24 of those participants being interviewed, the present study makes it possible to follow up with these teachers who are now practicing teachers.

Other Planetarium Software Programs for Future Studies

Although the use of *Starry Night*© has been shown to be a very useful approach to teaching inquiry-based lunar phase concepts, the cost can be prohibitive for teachers to use it in their classrooms. One possible way to encourage a wider use of planetarium software is to investigate the use of free planetarium software that is downloadable from the internet. Possible other software programs include: Stellarium (from Stellarium.org), Digital Universe (from the Hayden Planetarium), and Celestia (from sourceforge.net/projects/celestia). A systematic study should be conducted that would duplicate the present study with the use of various planetarium software packages. Each software package is slightly different in how lunar phases are rendered and how the operator interacts with the simulation. It would be important to investigate these software packages for ease of use and how applicable they are for simulating lunar phases in the night sky. These software packages perhaps might be easier to operate or perhaps much more difficult when compared to *Starry Night©*. The conceptual understanding of lunar phases should be measured by either forced-choice or interview methods while varying the software packages used. If a freely downloadable planetarium software package is equivalent, or perhaps superior, to *Starry Night©* then these other software packages could be recommended along with *Starry Night©.*

Demographic Future Studies

The LPCI and the Abridged-LPCI used in the present study also included questions regarding demographic data of the participants (see APPENDIX A). A future study could investigate this demographic data taken to see if there are any correlations of the demographic data with the student outcomes. In a related study that investigated factors affecting teacher self-efficacy among pre-service teachers, Cantrell et al. (2003) found that the effect of number of year of high school science was positively correlated and statistically significant. Demographic data collected in the present study using the LPCI included student reported self-confidence of answering the current assessment, college major, age range, gender, type of home town, ethnic background, highest level of math taken, self-reported math ability, and self-reported science ability. It would be interesting to compare the results of Cantrell et al. (2003) with the present study using self-reported science ability to see if there is a positive correlation with conceptual gains of lunar phases. Additionally, it would be interesting to see if there are differences in correlations between self-reported math and science abilities and the participants' highest level math class taken. Cantrell et al. (2003) found that males had a higher correlation between teaching self-efficacy and the amount of science taken. However, since all but 2 of the 69 participants were female in the present study, there would not be sufficient number of males in the present study to statistically investigate gender differences. Although, such a line of investigation would be important to conduct in the future because there appears to be gender differences in spatial reasoning, as discussed in Geary et al. (2000). Such spatial reasoning is necessary to build a conceptual model of lunar phases. In order to investigate these differences it would be necessary to expand the study into more traditional introductory astronomy classes where the demographic composition is more mixed.

Summary

This investigation lends further evidence for inquiry-based, conceptual lunar phase instruction for pre-service teachers. The interviews and student daily reflections were able to illuminate students replacing their initial mental models when confronted with disconfirming data. These initial mental models were often based on one or more alternative conceptions and were replaced with a revised model based on the data they collected that was scientifically accurate. Planetarium software can be an effective means for taking lunar observations. The use of the software can also greatly reduce the amount of time needed to gather moon phase data. There is statistically significant similarities between the using planetarium software in a whole-class demonstration mode and students taking their data while the operate laptop computers on their own.

The LPCI, Abridged-LPCI, and the semi-structured interview protocol can be useful instruments for measuring conceptual understanding of lunar phases. There is a statistically significantly correlated with the semi-structured interview for both pre-instruction and post-instruction, although this correlation lacks practical significance. However, the Abridged-LPCI can possibly be used in place of the interview in future studies that investigate large numbers of participants and where interviews are not practical especially if a sub-set of students is also interviewed to insure that the Abridged-LPCI is correlated to the interview for a specific situation. The use of the semi-structured interview protocol is still preferred when investigating small numbers of participants because interviews can better document the inquiry-based, conceptual change approach to science teaching.

The daily student reflections resulted in only limited correlations with gains in lunar phase conceptual understanding and these correlations were not practically significant. Nonetheless, the daily student reflections can be useful to gain insights into the teaching process, they are not generally recommended.

Computer simulations have made it possible for learners to simulate scientific observations that would either be too time-consuming or difficult to obtain in the real world. These computer simulations have been found to be either equivalent or superior to real world observations. The availability of computers in the classroom has made it possible to make regular use of computer simulations at all levels of science instruction, from elementary school to college. Students are therefore able to construct their own knowledge related to lunar phases by use of this constructivist approach. The continued use of inquiry-based, conceptual change instruction along with the use of planetarium software, such as *Starry Night©* is recommended for future instruction of all levels of learners, including elementary, middle, and high school levels in addition to pre-service teachers and practicing teachers. However, sometimes a simple approach such as using scientific models shows great promise as well. Therefore, the use of the scientific model activity is encouraged in the future all levels of instruction. Since a scientific conceptual understanding of lunar phases is so foundational, it is important that best practices are often revisited to insure that they are being recommended and used. As computer technology progresses, there will be future opportunities to investigate more immersive technologies such as virtual reality and manipulative 3-D imaging that will make it possible to more fully investigate the Earth-Moon-Sun spatial relationship.

REFERENCES

Abell, S., George, M., & Martini, M. (2002). The moon investigation: Instructional strategies for elementary science methods. *Journal of Science Teacher Education, 13* (2), 85-100.

Alexander, W. R. (2006). *Obstacles and challenges hindering sky observations in a college astronomy classroom.* Paper presented at the 207th meeting of the American Astronomical Society, Washington, DC, United States.

Alparslan, C., Tekkaya, C., & Geban, O. (2003) Using conceptual change instruction to improve learning. *Journal of Biology Education, 37* (3), 133-137.

American Association for the Advancement of Science. (1993). *Benchmarks for science literacy*. New York: Oxford Press.

Barab, S.A., Hay, K.E., Barnett, M., & Keating, T. (2000). Virtual Solar System Project: Building Understanding through Model Building. *Journal of Research in Science Teaching, 37*(7), 719-756.

Barab, S.A., Hay, K.E., Squire, K., Barnett, M., Schmidt, R., Karrigan, K., Yamagata-Lynch, L., & Johnson, C. (2000). Virtual Solar System Project: Learning through a technology-rich, inquiry-based, participatory learning environment. *Journal of Science Education and Technology, 9*(1), 7-25.

Barnett, M., & Moran, J. (2002). Addressing children's alternative frameworks of the moon's phases and eclipses. *International Journal of Science Education, 24* (8), 859-879.

Barnett, M., Yamagata-Lynch, L., Keating, T., Barab, S.A., & Hay, K.E. (2005). Using virtual reality computer models to support student understanding of astronomical concepts. *Journal of Computers in Mathematics and Science Teaching, 24*(4), 333-256.

Baxter, J. (1989). Children's understanding of familiar astronomical events. *International Journal of Science Education, 11*(special issue), 502-513.

Bell, R.L., Smetana, L., & Binns, I. (2005, October). Simplifying inquiry instruction: Assessing the inquiry level of classroom activities. *The Science Teacher, 72*(7), 30-33.

Bell, R.L. & Trundle, K.C. (2006). *Shooting for the Moon: Using a Computer Simulation to Facilitate Conceptual Change about Lunar Concepts.* A paper presented at the Annual Meeting of the National Association for Research in Science Teaching, San Francisco, CA, United States.

Bell, R., & Trundle, K. (2007). *Using technology to promote conceptual change: Phase 3 of the Research.* A paper presented at the Annual Meeting of the Association for Science Teacher Education, Clearwater Beach, FL, United States.

Bell, R.L., Bins, I.C., & Smetana, L.K. (2007). *Using technology to promote conceptual change in Secondary Earth Science Pupils' Understandings of Moon Phases*. A paper presented at the Annual Meeting of the Association for Science Teacher Education, Clearwater Beach, FL, United States.

Bell, R., & Trundle, K. (2008). The use of a computer simulation to promote scientific conceptions of moon phases. *Journal of Research in Science Teaching, 45*(3), 346-372.

Bisard, W. J., Aron, R. H., Francek, M. A., & Nelson, B. D. (1994). Assessing selected physical science and earth science misconceptions of middle school through university preservice teachers: Breaking the "misconception cycle". *Journal of College Science Teaching, 24,* 38-42.

Callison, P.L. & Wright, E.L. (1993, April). The effect of teaching strategies using models on presevice elementary teachers' conceptions about earth-sun-moon relationships. Paper presented at the annual meeting of the National Association for Research in Science Teaching, Atlanta, GA.

Cantrell, P., Young, S., & Moore, A. (2003). Factors affecting science teaching efficacy of preservice elementary teachers. *Journal of Science Teacher Education, 14* (3), 177-192.

Chang, C.Y. (2003). Teaching earth science: Should we implement teacher-directed or student-controlled CAI in the secondary classroom. *International Journal of Science Education, 25* (4), 427-438.

Crawford, T., Kelly, G.J., & Brown, C. (2000). Ways of knowing beyond facts and laws of science: An ethnographic investigation of student engagement in science practices. *Journal of Research in Science Teaching, 37*(3), 237-258.

Cropp, F. W., (1980). Student journals in physical geology. *Journal of Geological Education, 28*, 132-134.

deJong, T., & van Joolingen, W. R. (1998). Scientific Discovery Learning with Computer Simulations of Conceptual Domains. *Review of Educational Research, 68*(2), 179-201.

Deming, G. L. (2002). Results from the astronomy diagnostic test national project. *Astronomy Education Review, 1* (1), 52-57.

Edelson, D. (2000). Learning-for-use: A framework for the design of technology-supported inquiry activities. *Journal of Research in Science Teaching, 38*(3), 355-385.

Fan, X. 2001, Statistical Significance and Effect Size in Educational Research: Two Sides of a Coin, *Journal of Educational Research*, 94(5), 275.

Gazit, E., Yair, Y., & Chen, D. (2005). Emerging Conceptual Understanding of Complex Astronomical Phenomena by Using a Virtual Solar System. *Journal of Science Education and Technology, 14*(5/6), 459-470.

Geary, D. C., Saults, S. J., Liu, F., & Hoard, M. K. (2000). Sex differences in spatial cognition, computational fluency, and arithmetical reasoning. *Journal of Experimental Child Psychology*, 77, 337-353.

Hamilton, T.M. (2003). Computers in Science. *Journal of College Science Teaching, 32*(5), 322-326.

Hanson, J. A., Barnett, M., MaKinster, J. G., & Keating, T. (2004). The impact of three-dimensional computational modeling on student understanding of astronomical concepts: a quantitative analysis. *International Journal of Science Education, 26* (11), 1365-1378.

Hermann, R., & Lewis, B. F. (2003). Moon misconceptions: bringing pedagogical research of lunar phases into the classroom. *The Science Teacher*, *November*, 51-55.

Hobson, S.M., Trundle, K.C., & Saçkes, M. (2010). Using a Planetarium Software Program to Promote Conceptual Change with Young Children. *Journal of Science Education and Technology, 19*(2), 165-176.

Hogan, K., Nastasi, B.K., & Pressley, M. (2000). Discouse Patterns and Collaborative Scientific Reasoning in Peer and Teacher-Guided Discussions. *Cognition & Instruction, 17*(4), 379-432.

Hufnagel, B., Slater, T., Deming, G., Adams, J., Adrian, R., Brick, C., & Zeilik, M. (2000). Pre-course results from the astronomy diagnostic test. *Publications of the Astronomical Society of Australia, 17* (2), 152-155.

Hufnagel, B., (2002). Development of the astronomy diagnostic test. *Astronomy Education Review, 1* (1), 47-51.

Hyers, A. D., (2001). Predictable achievement patterns for student journals in introductory earth science courses. *Journal of Geography in Higher Education, 25* (1), 53-66.

Jones, B. L., Lynch, P. P., & Reesink, C. (1987). Children's conceptions of the earth, sun and moon. *International Journal of Science Education, 9*(1), 43-53.

Kangassalo, M. (1994). Children's independent exploration of a natural phenomenon by using a pictorial computer-based simulation. *Journal of Computing in Childhood Education, 5*(3/4), 285-297.

Kavanagh, C., Agan, L., & Sneider, C. (2005). A Learning about phases of the moon and eclipses: A guide for teachers and curriculum developers. *Astronomy Education Review, 4*(1), 19-52.

Keating, T., Barnett, M., Barab, S.A., & Hay, K.E. (2002). The Virtual Solar System Project: Developing Conceptual Understanding of Astronomical Concepts thought Building Three-Dimensional Computational Models. *Journal of Science Education and Technology, 11*(3), 261-275.

Keys, C.W., & Bryan, L.A. (2001). Co-constructing inquiry-based science with teachers: Essential research for lasting reform. *Journal of Research in Science Teaching, 38*(6), 631-645.

Kuethe, J. L. (1963). Science concepts: A study of "sophisticated" errors". *Science Education*, *47* (4), 361-364.

Lindell, R. S. (2001). Enhancing college students' understanding of lunar phases. *Unpublished Dissertation.* University of Nebraska-Lincoln.

Lindell, R. S. (2003). *Quantitatively investigating college students understanding of lunar phases*. Paper presented at the Enrico Fermi Summer School in Physics Education Research Conference Proceedings.

Lindell, R. S. (2004). *Measuring conceptual change in college students' understanding of lunar phases.* Proceedings of the 2004 Physics Education Research Conference, ed. by J. Marx , P. Heron, and S. Franklin. New York, NY: PERC Publishing. AIP Conference Proceedings; 2005, *790*(1), p 53-56.

Lindell, R. S. & Olsen, J.P. (2002). *Developing the lunar phases concept inventory.* Proceedings of the 2002 Physics Education Research Conference, ed. by S. Franklin, J. Marx and K. Cummings. New York, NY: PERC Publishing. http://piggy.rit.edu/franklin/PERC2002/Lindell.doc

Lindell, R. S. & Sommer, S. R. (2003). *Using the lunar phases concept inventory to investigate college students pre-instructional mental models of lunar phases*. Proceedings of the 2003 Physics Education Research Conference, ed. by S. Franklin, J. Marx and K. Cummings. New York, NY: PERC Publishing. AIP Conference Proceedings; 2004, *720*(1), p 73-76.

MacIntyre, B. (1999, December). Celestial Simulations. *The Science Teacher*, 36-39.

McDermott, L.C. (1996) *Physics by Inquiry (Volume 1).* New York: Wiley.

Miles, M.B., & Huberman, A.M. (1994). Qualitative data analysis: An expanded sourcebook, 2nd ed. London: Sage Publications.

Milne, C., & Otieno, T. (2007). Understanding engagement: Science demonstrations and emotional energy. *Science Education, 91*, 523-553.

Monaghan, J.M., & Clement, J. (1999). Use of a computer simulation to develop mental simulations for understanding relative motion concepts. *International Journal of Science Education, 21* (9), 921-944.

National Research Council. (1996). *National science education standards.* Washington, DC: National Academies Press.

National Research Council. (2013). *Next Generation Science Standards: For States, By States*. Washington, DC: National Academies Press.

Nehm, R.H., & Schonfeld, I.S. (2008). Measuring knowledge of natural selection: A comparison of the CINS, an open-response instrument, and an oral interview. *Journal of Research in Science Teaching, 45* (10), 1131-1160.

Olitsky, S. (2007). Promoting student engagement in Science: Interaction rituals and the pursuit of a community of practice. *Journal of Research in Science Teaching, 44*(1), 33-56.

Olsen, J.K. & Clough, M.P. (2001). Technology's Tendency to Undermine Serious Study: A Cautionary Note. *Clearing House, 75* (1), 8-13.

Parker, J. & Heywood, D. (1998). The earth and beyond: developing primary teachers' understanding of basic astronomical events. *International Journal of Science Education, 20* (5), 503-520.

Pasachoff, J. (2002). What should college students learn? Phases and season? Is less, more or is less less. *Astronomy Education Review, 1* (1), 124-130.

Ruiz-Primo, M.A. & Li, M. (2004). On the use of students' science notebooks as an assessment tool. *Studies in Educational Evaluation, 30*, 61-85.

Rutten, N., van Joolingen, W.R, & van der Veen, J.T. (2012). The Learning Effects of Computer Simulations in Science Education. *Computers & Education, 58* (1), 136-153.

Sadler, P. M (2000, February). Observational journals: An aid to sky watching. *Journal of College Science Teaching,* 245-254.

Schoon, K.J. (1992). Students' alternative conceptions of earth and space. *Journal of Geological Education*, *40*, 209-214.

Schoon, K.J. (1995). The origin and extent of alternative conceptions in the earth and space sciences: A survey of pre-service elementary teachers. *Journal of Elementary Science Education*, *7* (2), 27-46.

Schoon, K.J., & Boone, W. J. (1998). Self-efficacy and alternative conceptions of science of preservice elementary teachers. *Science Education*, *82*, 553-568.

Shadish, W.R., Cook, T.D., & Campbell, D.T. (2002). *Experimental and quasi-experimental designs for generalized causal inference.* Boston: Houghton Mifflin.

Smetana, L.K. (2008). The use of computer simulations in whole-class versus small-group settings. *Unpublished Dissertation.* University of Virginia.

Smetana, L.K., & Bell, R. L. (2012). Computer Simulations to Support Science Instruction and Learning: A critical review of the literature. *International Journal of Science Education, 34* (9), 1337-1370.

Smetana, L.K., & Bell, R. L. (2014). (in press)

Stahly, L. L., Krockover, G. H., & Shepardson, D. P. (1999). Third grade students' ideas about lunar phases. *Journal of Research in Science Teaching, 36* (2), 159-177.

Starver, J.R. (1998). Constructivism: Sound theory for explicating the practice of science and science teaching. *Journal of Research in Science Teaching*, *35* (5), 501-520.

Stoker, H. W., King, F. J., & Foster, B. F. (1981). The class as an experimental unit. *Evaluation Review*, *5* (6), 851-859.

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N.K. Denzin & Y.S. Lincoln (Eds.), Handbook of qualitative research (pp.273-285). London: Sage Publications.

Suzuki, M. (2003). Conversations about the moon with prospective teachers in Japan. *Science Education*, *87*, 892-910.

Tessier, J., (2007). Small-Group Peer Teaching in an Introductory Biology Classroom. *Journal of College Science Teaching, 36* (4), 64-69.

Tien, L.T, Roth, V., Kampmeier, J.A., (2002). Implementation of a Peer-Led Team Instructional Approach in an Undergraduate Organic Chemistry Course. *Journal of Research in Science Teaching, 39*(7), 606-632.

Trumper, R. (2000). University students' conceptions of basic astronomy concepts. *Physics Education, 35* (1), 9-15.

Trundle, K. C., Atwood, R. K., & Christopher, J. E. (2002). Preservice elementary teachers' conceptions of moon phases before and after instruction. *Journal of Research in Science Teaching, 39* (7), 633-658.

Trundle, K. C., Atwood, R. K., & Christopher, J. E. (2006). Preservice elementary teachers' knowledge of observable moon phases and pattern of changes in phases. *Journal of Science Education, 17* (2), 87-101.

Trundle, K. C., Atwood, R. K., & Christopher, J. E. (2007a). Fourth-grade elementary students' conceptions of standards-based lunar concepts. *International Journal of Science Education, 29* (5), 595-616.

Trundle, K. C., Atwood, R. K., & Christopher, J. E. (2007b). A longitudinal study of conceptual change: Preservice elementary teachers'conceptions of moon phases. *Journal of Research in Science Teaching, 44* (2), 303-326.

Trundle, K. C., Atwood, R. K., Christopher, J. E., & M. Saçkes (2010). The Effect of Guided Inquiry-Based Instruction on Middle School Students' Understanding of Lunar Concepts. *Research in Science Education, 40* (3), 451-478.

Trundle, K.C. & Bell, R.L. (2003). Technology reviews: Using planetarium software to teach standards-based lunar concepts. *School Science and Mathematics, 103* (8), 397-401.

Trundle, K.C. & Bell, R.L. (2010). The Use of a Computer Simulation to Promote Conceptual Change: A quasi-experimental study. *Computers & Education, 54* (4), 1078-1088.

Trundle, K.C. & Troland, T. H. (2005, October). The moon in children's literature. *Science and Children,* 40-43.

Vosniadou, S. (1994). Capturing and modeling the process of conceptual change. *Learning and Instruction, 4,* 45-69.

Vosniadou, S., & Brewer, W. F. (1994). Mental models of the day/night cycle. *Cognitive Science, 18*, 123-183.

Vosniadou, S., Skopeliti, I., & Ikospentaki, K. (2004). Modes of knowing and ways of reasoning in elementary astronomy. *Cognitive Development, 19*, 203-222.

White, B.Y. (1993). ThinkerTools: Causal models, conceptual change, and Science Education. *Cognition & Instruction, 10*(1), 1-100.

Windschitl, M. & Andre, T. (1998). Using computer simulations to enhance conceptual change: The roles of constructivist instruction and student epistemological beliefs. *Journal of Research in Science Teaching, 35* (2), 145-160.

Zeilik, M. (2003). Birth of the Astronomy diagnostic test: Prototest evolution. *Astronomy Education Review, 1* (2), 46-52.

Zeilik, M., & Bisard, W. (2000). Conceptual change in introductory-level astronomy courses: Tracking misconceptions to reveal which-and how much-concepts change. *Journal of College Science Teaching, 29* (4), 229-232.

Zeilik, M., & Morris, V. (2003). An examination of misconceptions in an astronomy course for science, mathematics, and engineering majors. *Astronomy Education Review, 2* (1).

Zeilik, M., Schau, C., & Mattern, N. (1998). Misconceptions and their change in university-level astronomy courses. *The Physics Teacher*, 36, 104-107.

Zeilik, M., Schau, C., & Mattern, N. (1999). Conceptual astronomy. II. Replicating conceptual gains, probing attitude changes across three semesters. *American Journal of Physics, 67* (10), 923-927.

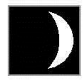
APPENDIX A:

LUNAR PHASE CONCEPT INVENTORY (LPCI)

Lunar Phases Concept Inventory

This concept inventory is designed to assess your knowledge of the Lunar Phases. Please choose the best answer for each question and record it on the bubble sheet provided. Answer each question to the best of your ability. Do not mark on this concept inventory. Please return this inventory with your bubble sheet.

1. Immediately after Sunset, you observe a Moon with the following shape.



The Moon would:
  a. be just rising.
  b. be at its highest position in the sky.
  c. be just setting
  d. not see because it has yet to rise.
  e. not see because it has already set.
  f. be anywhere in the sky.

2. The Moon orbits around the Earth; approximately how long does it take to complete one orbit?

a. Less than one day b. One day c. One week d. Two weeks
e. One month f. 6 months g. One year h. More than one year

3. The Moon orbits around the Earth; in which direction does it orbit if observed from a point directly above the Earth's North Pole?

a. Clockwise b. Counter Clockwise c. Either Direction

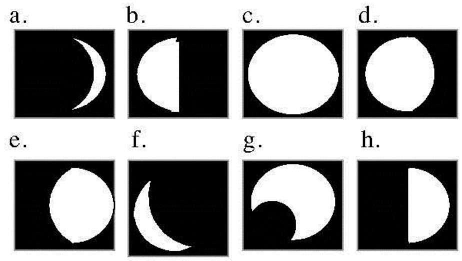
4. A *New Moon* occurs when no lighted portion of the moon is visible to an observer on Earth This occurs because
  a. an object completely blocks the Moon.
  b. the Moon is completely covered by the shadow of the Sun.
  c. the Moon is completely covered by the shadow of the Earth.
  d. the moon is between the Earth and the Sun.
  e. Both a and d.
  f. Both b and d
  g. Both c and d
  h. None of the above

5. How often do *New Moons* occur?

a. Every day/ night b. Once a week c. Once every two week d. Once a month e. Once a year

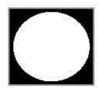
1 of 7

6. When the Moon appears to completely cover the Sun (an eclipse), the Moon must appear to have which shape/ phase?

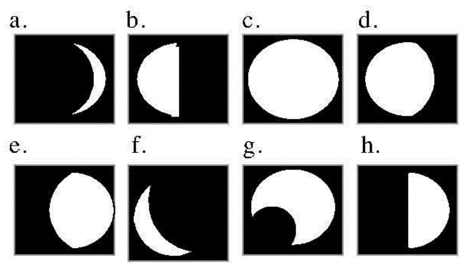


i. The moon will appear completely dark (New Moon).
j. Can be any shape/ phase.

7. An Australian friend observed a Moon with the following shape tonight,



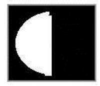
What shape would the Moon have if you observed at your current location later tonight?



i. none of the above
j. any of the above

8. The time it takes the Moon to orbit the Earth and the time it takes the Moon to complete a cycle of phases have which of the following relationships?
   a. The Moon completes its cycle of phases in much less time than it takes to orbit the Earth.
   b. The Moon completes its cycle of phases in approximately the same time as takes to orbit the Earth.
   c. The Moon completes its cycle of phases in much more time than it takes to orbit the Earth.
   d. There is no relationship between the cycle of phases and the time it takes the Moon to orbit the earth.

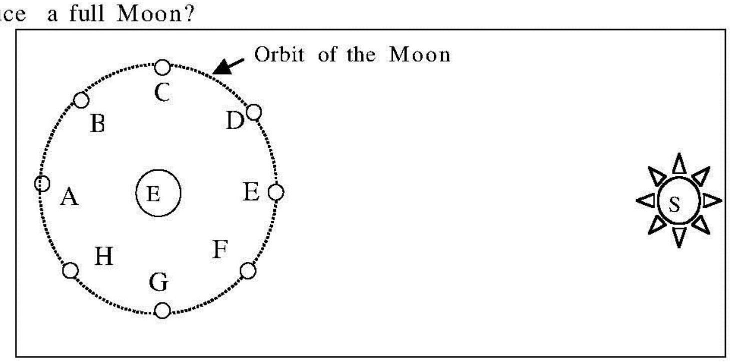
9. You observe a Moon with the following shape just setting.



What time is it?
a. Dawn
b. Midway between Dawn and Noon
c. Noon
d. Midway between Noon and Sunset
e. Sunset
f. Midway between Sunset and Midnight
g. Midnight
h. Midway between Midnight and Dawn
i. Any time of the night
j. Any time of the day or night

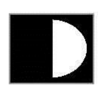
10. Which direction do you look to observe the moon in question 9 set?
a. North b. East c. South d. West e. Any direction

11. If you could look down on the Earth/ Moon/ Sun system from a point in deep space located above the Earth's North Pole, you could observe the following alignments. Which Earth-Moon–Sun geometry would produce a full Moon?



i. none of the above j. any of the above

12. You observe a Moon with the following shape.

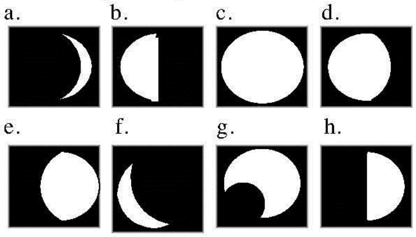


How long until the Moon looks this way again?

a. 1 day later	b. 1week later	c. 2 weeks later	d. 3 weeks later
e. 1 month later	f. 6 months later	g. 1 year later	h. None of the above

13.If you were to observe the Moon in question 12, in London, England what shape would you expect the Moon to have?

a. b. c. d.



e. f. g. h.

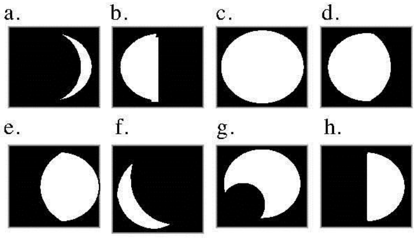
i. none of the above j. any of the above

14.As the Sun sets, you observe a Moon just rising. What direction would you look to observe the Moon?

a. North b. East c. South d. West e. Any direction

15.What shape would the Moon in question 14 have?

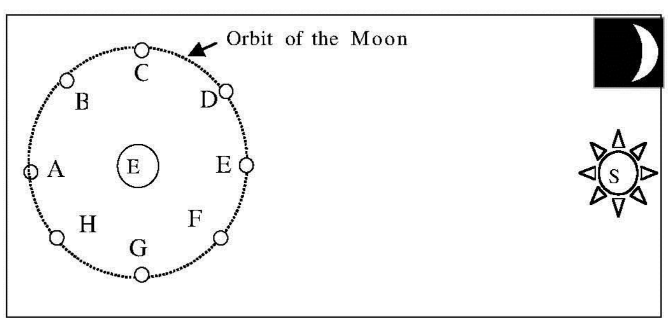
a. b. c. d.



e. f. g. h.

i. none of the above j. any of the above

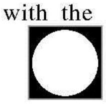
16.If you could look down on the Earth/Moon/Sun system from a point in space located above the Earth's North Pole, you could observe the following alignments. Which Earth-Moon–Sun geometry would produce a Moon with a shape shown below?



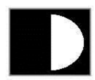
i. none of the above j. any of the above

Lunar Phases Concept Inventory

17. You observe the moon with the following shape and phase tonight,



At some later time, you observe the moon again and it has the following shape and phase.

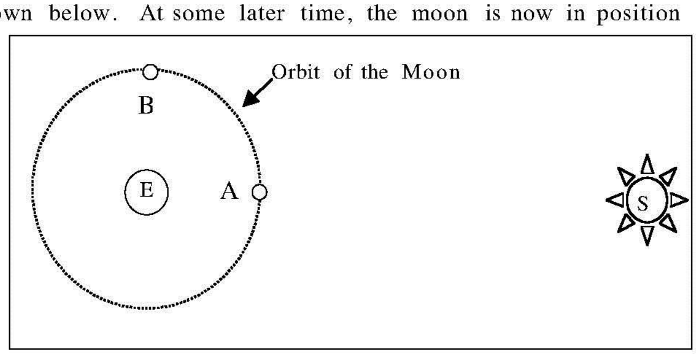


How long between your two observations?

a. 1 hour	b. 3 hours	c. 6 hours	d. 1/2 day	e. One day
f. 1 week	g. 2 Weeks	h. 3 Weeks	i. 1 month	j. More than 1 month

18. What caused the moon to appear different in your two observations in question 17?
  - a. an object is now between the Moon and the Earth
  - b. the Moon is now covered by the shadow of the Sun
  - c. the Moon is now covered by the shadow of the Earth.
  - d. the Moon's position relative to the Earth has changed.
  - e. Both a and d.
  - f. Both b and d
  - g. Both c and d
  - h. None of the above

19. If you could look down on the Earth/Moon/Sun system from a point in space located above the Earth's North Pole, you would observe that the Moon orbits around the Earth. At one point in time it is in position A, as shown below. At some later time, the moon is now in position B.



How much time passed between these two observations?

a. 1 hour	b. 3 hours	c. 6 hours	d. 1/2 day	e. One day
f. 1 week	g. 2 Weeks	h. 3 Weeks	i. 1 month	j. More than 1 month

20.Which direction did the moon travel around the Earth?
a. Clockwise b. Counter Clockwise c. Either Direction

21.In general, how confident are you that your answers to this survey are correct?
a. Not at all confident (just guessing)
b. Not very confident
c. Not Sure
d. Confident
e. Very Confident

22.What is your college major (or current area of interest if undecided)?
a. Business
b. Education
c. Humanities, Social Science or the Arts
d. Science, Engineering, Agriculture or Architecture
e. Other

23.What is your age?
a. 0-20 years old
b. 21-23 years old
c. 24-30 years old
d. 31 or older
e. Decline to answer

24.What is your gender?
a. Female
b. Male
c. Decline to answer

25.Which best describes your home community (where you attended high school)?
a. Rural
b. Small Town
c. Suburban
d. Urban
e. Not in USA

26.Which best describes your ethnic background?
a. African American
b. Asian-American
c. Native American
d. Hispanic-American
e. African (not American)
f. Asian (not American)
g. White, non Hispanic
h. Multicultural
i. None of the Above
j. Decline to answer

27. What was the highest level math class you completed prior to taking this course?
   a. Algebra
   b. Trigonometry
   c. Geometry
   d. Pre-Calculus
   e. Calculus

28. How good at math are you?
   a. Very Poor
   b. Poor
   c. Average
   d. Good
   e. Very Good

29. How good at science are you?
   a. Very Poor
   b. Poor
   c. Average
   d. Good
   e. Very Good

APPENDIX B:

Instructional Treatments

For treatment A, the instructor will operate the *Starry Night©* software in demonstration mode for the entire class and the students will take data in groups of two. For treatment B, the students will be operating the software in groups of two and taking data. Observations will start the class meeting that immediately follows a new moon, so that the observations will be related to what is currently going on in the sky. The participants will keep track of their observations on two different types of data sheets. The first data sheet records the position of the moon with respect to the horizon looking west and is shown below (Figure 1). The participants will change their data sheets and use the one viewing east (Figure 2) after the first week of observations because the moon shifts from west to south, then towards the east. After the second week, another copy of the chart viewing east will be used, but this time the observations will be made approximately one hour later each subsequent night.

Lunar Phase (position) Observing Record Chart-West

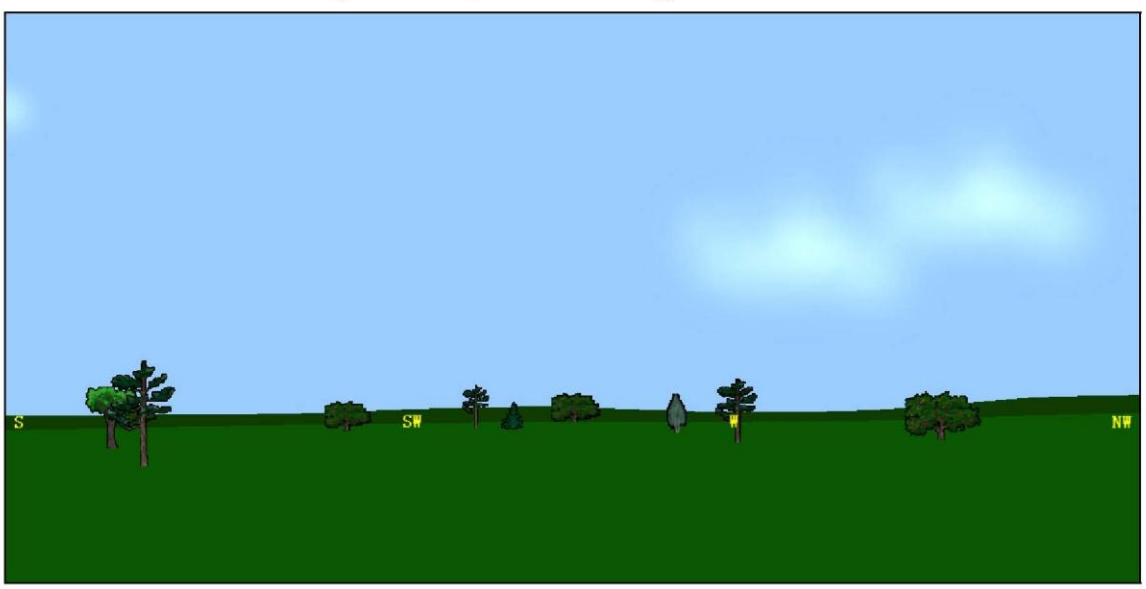


Figure 1. Data sheet for position of lunar phase observations looking west.

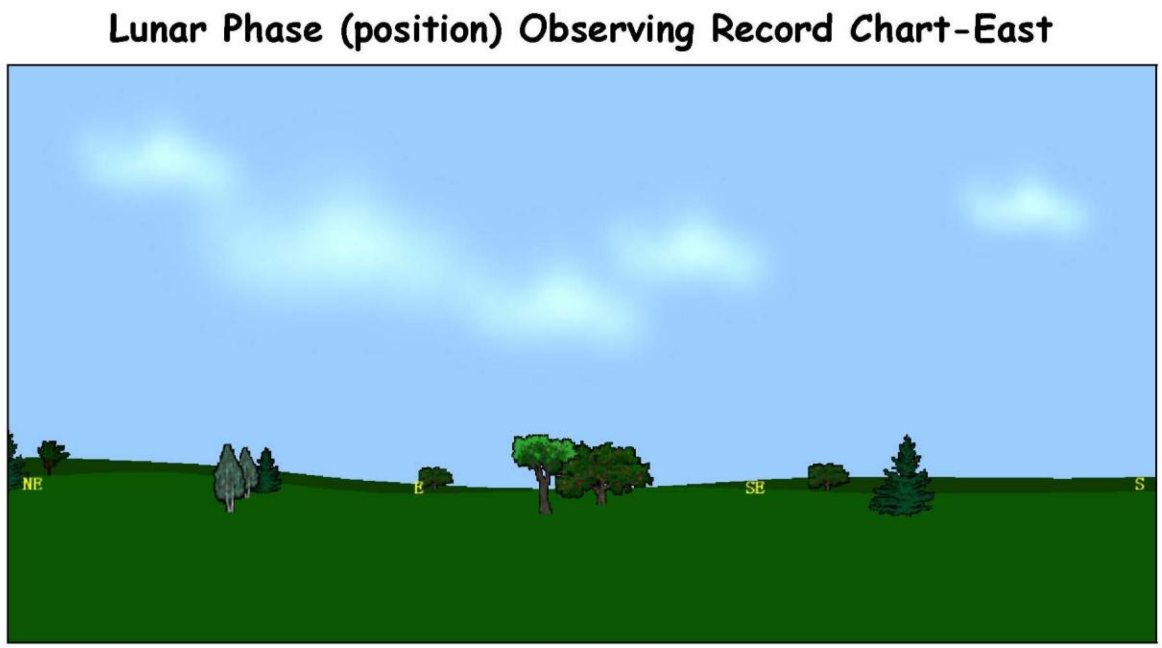


Figure 2. Data sheet for position of lunar phase observations looking east.

Participants will keep a record of their observations by recording the shape, direction in the sky, and the date and time of the observation using the data sheet (Figure 3) shown below. Data will be gathered only up to the current date for each class meeting. Each group will make observations during the waxing phases, from new moon to full moon and then the waning phases up to the 3rd quarter moon. In addition to the daily sketches that students will make, they will also write a daily reflection for each day they make observations. These student reflections will include the following elements:

- Drawings made during instruction
- Description of the data gathered
- Possible explanation of the data gathered
- Predictions related to the cause of the lunar phases
- Questions for next class
- Any major surprises from the day

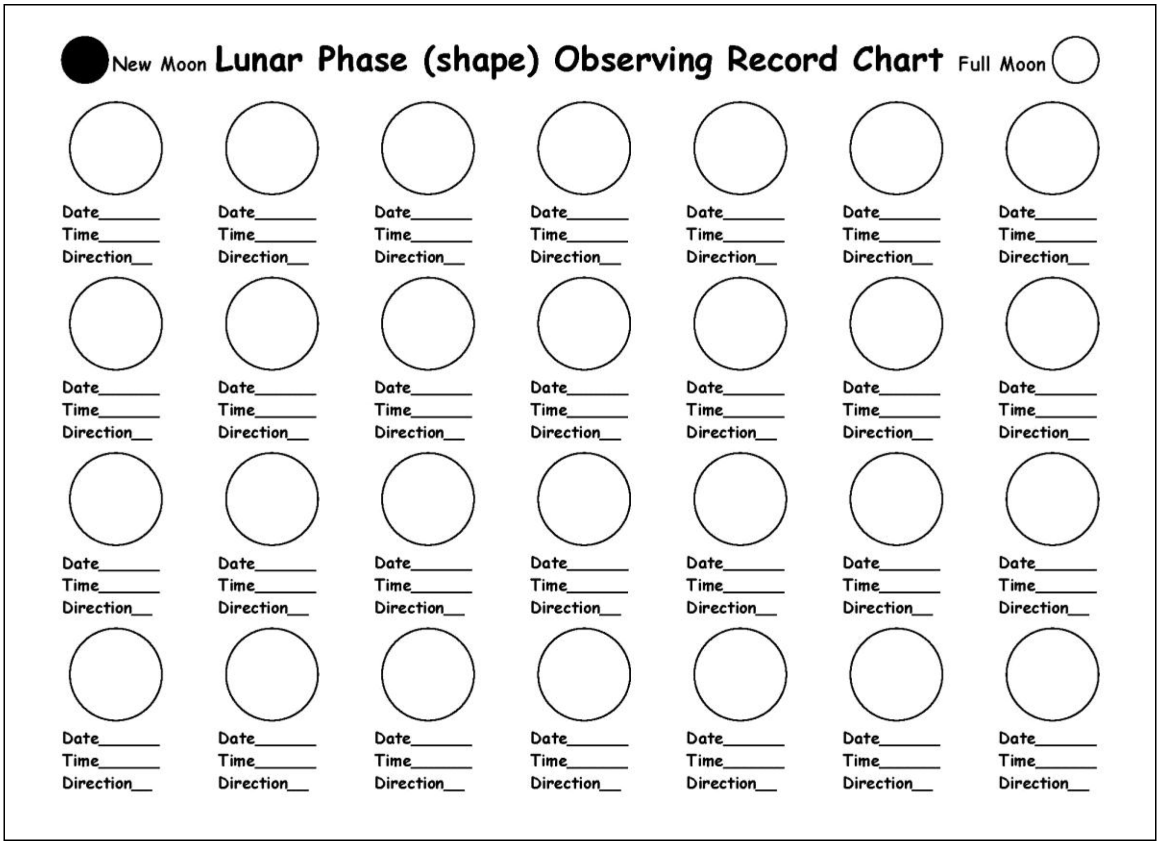


Figure 3. Data sheet for date, time, and description of position for lunar phase observations.

After the observations are complete, the participants will devote one class period (1.5 hours) to the analysis of their observations. This is a critical part of the treatment which will include five separate sections that address the following topics: 1) describing patterns in the observations such as shape and sequence, 2) applying new concepts and terminology, 3) making a prediction of the length of the overall lunar cycle as well as how the shape changes in a single day, 4) making three-dimensional models to simulate the positions of the moon at various phases and to predict the rising and setting time of the moon at various phases, 5) using the three-dimensional models to simulate the positions of lunar and solar eclipses to directly address the issue of role of shadows in observing the moon. The goal of this treatment is to have students apply their knowledge of lunar phases to construct physical models that show how lunar phases, thus they are applying their knowledge.

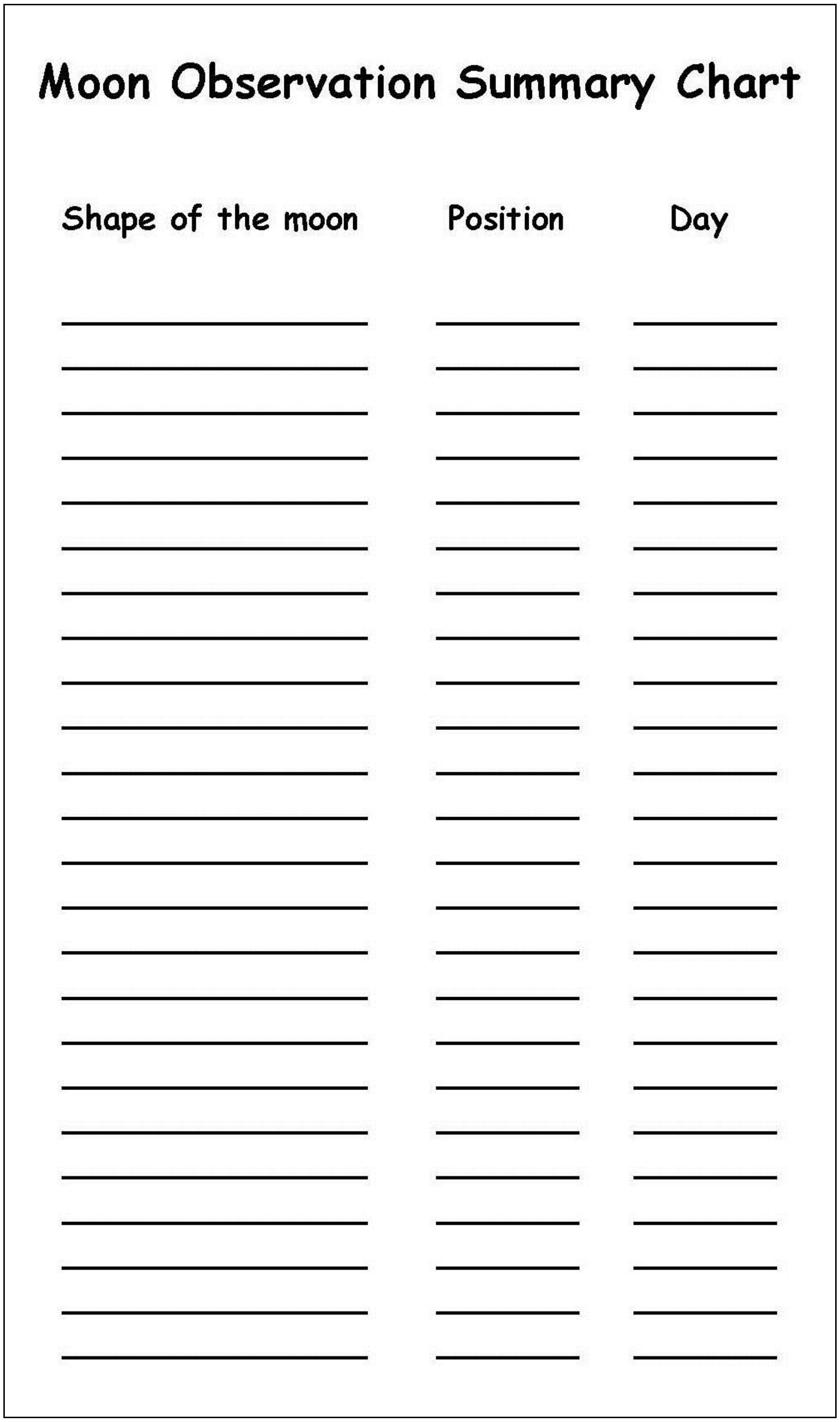


Figure 4. Moon observation summary chart

The students will make a prediction of how the lunar phases will change once the Styrofoam ball representing the moon passes the position for 3rd quarter moon. They will make a sketch in their reflections describing their predictions along with an explanation of their model on which they have based their predictions. They will then test their predictions by using the 3-D models as well as checking their predictions using *Starry Night©*. They will report in their reflections how their predictions performed and any discrepancies they encountered. They will then sketch a revised model based on what they learned from testing their mental model of lunar phases.

The students will then work in groups on a series of questions to extend their knowledge related to the time of day certain lunar phases are visible and location in the sky. These extension questions are listed below:

i. Are the sun and the moon ever in the sky at the same time?
ii. Where is the sun relative to the moon when the moon is full?
iii. Where is the sun relative to the moon when the moon is first quarter?
iv. Where is the sun relative to the moon when the moon is third quarter?

APPENDIX C:

SCORING RUBRIC FOR INTERVIEW PROTOCOL (Trundle et al. (2010))

Scoring Rubric for Interview Protocol (Conceptual Understanding)

Scientific:		Participant's conceptual understanding exhibits all element of scientific understanding without exhibiting alternative conception.
	10 Points	Includes all elements of scientific understanding.
Scientific fragment:		Participant's conceptual understanding does not exhibit an alternative mental model, but fails to include all elements of scientific understanding.
	9 Points	Missing one element of scientific understanding.
	8 Points	Missing two elements of scientific understanding.
	7 Points	Missing three elements of scientific understanding.
Scientific with alternative Fragment:		Participant exhibit all <u>four</u> elements of scientific understanding along with an alternative mental model.
	6 Points	Includes all elements of scientific understanding with an alternative mental model.
Alternative with Scientific fragments:		Participant's conceptual understanding exhibits an alternative mental model, but also includes some elements of scientific understanding.
	5 Points	Includes an alternative mental model, but also contains three elements of scientific understanding.
	4 Points	Includes an alternative mental model, but also contains two elements of scientific understanding.
	3 Points	Includes an alternative mental model, but also contains one element of scientific understanding.
Alternative:		Participant' conceptual understanding exhibits no elements of scientific understanding and includes a single mental model.
	2 Points	Includes a single alternative mental model without any elements of scientific understanding.
Alternative fragments:		Participant' conceptual understanding exhibits two or more alternative mental models. Conceptual understanding may or may not exhibit some elements of scientific understanding.
	1 Points	Includes two or more alternative mental models.
No conceptual Understanding:		Participant exhibits no conceptual understanding.
	0 Points	Participant exhibits no conceptual understanding.