THE EFFECTS OF COMPUTER-SUPPORTED INQUIRY-BASED LEARNING METHODS AND PEER INTERACTION ON LEARNING STELLAR PARALLAX

A Dissertation

presented to

the Faculty of the Graduate School

at the University of Missouri-Columbia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

---

by

LANIKA RUZHITSKAYA

Dr. James Laffey, Dissertation Co-advisor
Dr. Angela Speck, Dissertation Co-advisor

JULY 20011

The undersigned, appointed by the dean of the Graduate School, have examined the entitled

THE EFFECTS OF COMPUTER-SUPPORTED INQUIRY-BASED LEARNING METHODS AND PEER INTERACTION ON LEARNING STELLAR PARALLAX

Presented by Lanika Ruzhitskaya

A candidate for the degree of Doctor of Philosophy

And hereby certify that, in their opinion, it is worthy of acceptance.

Professor James Laffey

Professor Angela Speck

Professor David Jonassen

Professor John Wedman

To my father, who did not live to see the fulfillment of his dream…

ACKNOWLEDGEMENTS

There are many people who I want to thank for their encouragement and support on my way to receiving a doctoral degree. Without their advice and friendship this journey would have never begun and would certainly never have come to the end.

I would like to express my deepest gratitude to my both advisors: Dr. Jim Laffey and Dr. Angela Speck. Professor Laffey, who served as my advisor and the committee chair, transformed me from just a developer of computer applications to a researcher. He taught me not only how to design a study and chose the right sample size, but also how to look for valuable research questions and to find the crux for a meaningful investigation. I learned a great deal from your guidance, thank you.

It is with Dr. Laffey's support of my interest in astronomy my collaborative work with Dr. Speck led me to my career choice, and to my development as a researcher in the field of astronomy education. Professor Speck made me to believe in myself first as a student in her astronomy courses and then as an astronomy instructor and a researcher. She encouraged my ideas of developing educational tools, implementing and finally testing them in her classroom. It was with her interest in my work and encouragement that we created *Stellar Properties Laboratory* that consequently led us to creation of *SPIRUT* – my dissertation work. Thank you, professor, with all my heart.

I would like to extend my grateful appreciation to my committee members Dr. John Wedman and Dr. David Jonassen for providing me with your feedback, pointing out flaws in the study and giving me your suggestions on how to improve my research throughout my work on the thesis.

I would like to thank my colleagues who helped me to develop and improve my assessment instrument and worked with me on taxonomic level of questions: Rica S. French, a professor of Astronomy Collaboration of Astronomy Teaching Scholars (*CATS*), Benjamin Mendelsohn, a professor at West Valley College in Saratoga, California, and Dr. Seth Hornstein, a senior instructor in the department of Astrophysical and Planetary Sciences at University of Colorado, Boulder.

I would like to extend my sincere gratitude to my friends who helped me with my thesis's data analysis: Laura Hosmer, Lucas Miller, Brandon Hester, and Woodrow Petrović. And to the staff of Allen Institute Gary Westergren, Aaron Moss, Christian Cepel, and Neeley Current who helped me in setting up the study. And I would like to thank my friends Piyanan Nuankhieo, Musa Olaka, Borchuluun Yadamsuren, Camille Dickson-Deane, Ashkan Shafiee, David Arrant, John and Heather Gaddy whose support was most appreciated and valuable on so many occasions during the last several years of my life.

Finally, my dearest appreciation goes to my family: my mother, my brother, my sister-in-law and my nephew, who patiently supported me from the other side of the world, and have never doubted the path I have chosen.

TABLE OF CONTENTS

CHAPTERS

CHAPTER 5. DISCUSSION AND CONCLUSION ……………………………..208-246

APPENDICES

LIST OF ILLUSTRATIONS

LIST OF TABLES

Table Page

THE EFFECTS OF COMPUTER-SUPPORTED INQUIRY-BASED LEARNING METHODS AND PEER INTERACTION ON LEARNING STELLAR PARALLAX

Lanika Ruzhitskaya

Dr. Jim Laffey, Dissertation Co-advisor

Dr. Angela Speck, Dissertation Co-advisor

ABSTRACT

The presented research study investigated the effects of computer-supported inquiry-based learning and peer interaction methods on effectiveness of learning a scientific concept. The stellar parallax concept was selected as a basic, and yet important in astronomy, scientific construct, which is based on a straightforward relationship of several components presented in a simple mathematical equation: $d = 1/p$. The simplicity of the concept allowed the researchers to explore how the learners construct their conceptual knowledge, build mathematical skills and transfer their knowledge beyond the learning settings. A computer-based tutorial *Stellar Parallax Interactive Restricted and Unrestricted Tutorial* (*SPIRUT*) was developed for this study, and was designed to aid students' knowledge construction of the concept either in a learner-controlled or a program-controlled mode. The first investigated method in the study was enhancing engagement by the means of scaffolding for inquiry, which included scripted prompts and called for students' predictions and reflections while working in the learner-controlled or the computer-controlled version of *SPIRUT*. A second form of enhancing engagement was through peers working cooperatively during the learning activities. The students' level of understanding of the concept was measured by (1) the number of correct answers on a conceptual test with (2) several questions that require knowledge transfer to unfamiliar situations and (3) their ability to calculate the stellar parallax and find distances to stars.

The study was conducted in the University of Missouri among 199 non-science major students enrolled in an introductory astronomy course in the fall semester 2010. The participants were divided into two main groups: one was working with *SPIRUT* and another group was a control group and utilized a paper-based tutorial. The *SPIRUT* group was further divided into the learner-controlled and the program-controlled subgroups. Students' learning achievements were measured by two post- tests and compared to the students' results on a pre-test. The first post-test was administered right after the treatment with aim to measure the immediate effect of the treatment. The second post-test was administered eight weeks later and was aimed to elicit how much of the constructed knowledge students retained after the treatment.

Results of the study revealed that students who learned the concept with *SPIRUT* constructed greater conceptual knowledge and were able to better transfer it to another situation while their mathematical skills were equally improved as those students who worked with the paper-based tutorial. It was also evident that there was no difference between students' performances after their engagement with the learner-controlled or with the program-controlled version of *SPIRUT*. It was also found that students who worked independently constructed slightly greater knowledge than students who worked with peers. Albeit, there was no significant difference found of retention of knowledge after any type of treatment.

CHAPTER 1

INTRODUCTION

I. Introduction

The goal of the study is to investigate the effects of computer-supported inquiry-based learning and peer interaction methods on learning a scientific concept.

Emphasis on use of technologies to support learning in a classroom (Herman, 1996) started a chain reaction that resulted in creating a number of computer-based learning tools: from simple visualizations to complex high-fidelity three-dimensional virtual learning environments (Marshall, 1998; Ruzhitskaya & Speck, 2008). It has been noted by education researchers that technology-based learning can help learners' to improve their reading ability (Solomon, Globerson & Guterman, 1989), decision making skills (Berson, 1996), problem-solving skills (Jonassen, 1996; Jonassen, Prevish, Christy & Stavurlaki, 1999), data processing and computational skills (Suh & Moeyr-Packenham, 2007), and foster conceptual change (Zacharia & Anderson, 2003). In addition, computer-based environments support the learning process by scaling down real-world phenomenon, simplifying it, and/or accentuate its particular features to help students to concentrate on the desired concept (van Joolingen & de Jong, 1991), provide necessary tools for data visualization and analysis (Ozmen, 2007), and facilitate collaboration (King, 2002).

Engaging students in a scientific inquiry is a well-established method of facilitating students' learning through their direct participation in the science process:

from recognizing facts and formulating predictions to forming scientific models and theories (Chang, Sung, & Lee, 2003; de Jong, 2006b). A number of of investigations looked at efficacy of a range of computer-supported scaffolding methods designed for supporting computer-based inquiry learning and found them beneficial for learners' conceptual change (Hmelo-Silver, 2006; Hmelo-Silver, Duncan, & Chinn, 2007). Additionally, computer-based environments create authentic settings in which learners engaged in inquiry tasks of a scientific process and supported with scaffolding computer-based tools (van Joolingen, de Jong, & Dimitrakopoulout, 2007). In addition, a number of computer environments are designed to facilitate students learning through collaboration with peers. A number of studies were conducted to find out more when peer interaction is beneficial and when it becomes an impediment to learning (Ardac & Sezen, 2002; Crooks, Klein, Savenye, & Leader 1998; Hooper, Temiyakarn, & Williams 1991; King, 1998; Singhanayok & Hooper, 1998). In addition, the role of learner-controlled and computer-controlled environments is not as clear (Gay, 1986; Hannafin, 1984; Hartley, 1985; Kopcha & Sullivan, 2007; Lawless & Brown, 1997; McNeil & Nelson, 1991; Penland, 1979; Ross & Rakow, 1981; Williams, 1993; Young, 1996).

The main goal of this study is to investigate further the role of computer-supported tools and the role of peer interaction on students' learning of a scientific concept.

The study uses a computer-based tutorial *Stellar Parallax Interactive Restricted and Unrestricted Tutorial* (*SPIRUT*) designed to aid students' knowledge construction of the stellar parallax which is the first step in students' learning of stellar properties - one of the most important concepts in astronomy courses. *SPIRUT* was also designed to assist students in their performance on the *Stellar Properties Laboratory* (*SPLab*) project, which combines students' exploratory work of several stellar properties (Ruzhitskaya & Speck, 2007).

The study examines how methods such as inquiry-based computer-assisted instruction in the form of an interactive tutorial and peer interaction enhances students' engagement in the learning process so as to build their conceptual understanding and facilitate transfer of knowledge. The first method of enhancing engagement is scaffolding for inquiry, which includes scripted prompts and calls for students' predictions and reflections while undertaking the tutorial. A second form of enhancing engagement is through having a peer who can work interactively during the learning activity. The students' level of understanding is measured by (1) the number of correct answers on a conceptual test with (2) several questions that require knowledge transfer to unfamiliar situations and (3) their ability to calculate the stellar parallax.

The test of the computer-supported inquiry-based approach compares the achievements of a group of students using scaffolding (for making predictions and reflecting on observed changes as a result of changing simulated variables related to finding a stellar parallax angle) during a computer-based tutorial with the achievement of another group of students using a paper-based tutorial, *Lecture-Tutorials* (*LTs*) (Prather, Slater, Adams, & Brissenden, 2008). A detailed description of the program is included starting in Chapter 3, page 74.

The other goal of the proposed study is to investigate the influence of a social component – namely the collaboration between peers - on students' learning outcomes. To test the effect of collaboration, achievements for students working individually through the simulations are compared to the achievements of students working in pairs with their peers.

To assess the learners' achievement a *Stellar Parallax Assessment* (*SPA*) instrument (see Appendix C) has been developed; this instrument was developed using items from several standardized tests and several questions were developed by the investigator in collaboration with three astronomy and physics instructors.

The study was conducted with 199 non-science major students enrolled in an introductory astronomy course at the University of Missouri. The discussed study is based on one of the many topics taught in this course: an introduction to the basic properties of stars. The study utilizes the *SPLab* project, which includes a computer-based educational simulation with a guided introduction to stellar parallax, *SPIRUT*, as well as a paper-based manual guidance to the laboratory work. *SPIRUT* encompasses several types of scaffolding for scientific inquiry: prediction, observations/experiments, and reflection. The tutorial also provides learners with feedback on their answers to multiple-choice questions.

Two forms of *SPIRUT* have been developed: learner-control and program-control versions. The first version gives its users a full control of the interactive environment by allowing users to select exercises they want to work at, go forward and back, skip questions or work on different parts out of sequence. Another version of *SPIRUT* is a program-controlled environment in which students are taken through a sequential learning process selected for them by the tutorial based on students' performances on presented tasks. A detailed description of the program is included starting in Chapter 3, page 75.

II. Background

II.I. Astronomy Education

The National Science Education Standards (National Research Council [NRC], 1996) tells educators to use new and more constructivist approaches in teaching science. The standards stress the development of conceptual understanding of scientific constructs. The NRC places emphasis on the students' ability to provide explanations, generate predictions, and create arguments from observed evidence (Edelson, 2001).

According to the American Institute of Physics, approximately 250,000 students enroll into introductory astronomy courses in the United States each year (Fraknoi, 2001; Prather, Rudolph, Brissenden, & Schlingman, 2009). While some students have a genuine interest in astronomy, most of the students are non-science major students who take the course in order to fulfill their institutions' general education requirements for science education credits (Bailey, Prather, & Slater, 2003). Many of the pupils do not possess deep knowledge of mathematics and physics. Moreover, this introductory astronomy course may very-well be the only science course these students will ever take (Prather, et al.). Therefore, the introductory astronomy courses are designed to help students gain conceptual understanding of selected topics in astronomy, stimulate their interest in science, and help them to develop their critical thinking skills (Marschall, 2000; Slater & Adams, 2000).

It is commonly understood that astronomy is one of the most difficult subjects to learn. In addition to reasoning skills and the application of physics and mathematics, astronomy requires understanding of complex concepts that often involve other scientific disciplines such as chemistry, geology, biology, to name a few (Marschall, et al.). As a result, it is not surprising that many people are intimidated by this discipline - a complicated knot of sciences.

Astronomy is also a challenging subject to teach. It involves a vast amount of information ranging from the first Hydrogen atom coming into existence after the Big Bang all the way to the effects of the absorption of infrared radiation by the Earth's atmosphere on the life on our planet. Consequently, one of the problems that students encounter in an introductory astronomy course is the large amount of concepts squeezed into a relatively short period of time (Sadler, 1996). As a result, students often have to move on to a harder concept without mastering understanding of the simpler one. This is dangerous and a sure path for creating misconceptions. In his study, Sadler pointed out that the concepts cannot be taught randomly but rather should be chosen "on the basis of their structural relationship to learning other concepts" (p.54). He proposed the existence of a conceptual hierarchy in the process of learning astronomy. While preconceptions - a certain basic knowledge about a subject - is a prerequisite for understanding more difficult topics (Sadler, et al), without a solid basis for advancing the knowledge, it can become a misconception. In his empirical study based on *Project STAR*, Sadler found that "an attempt to cover too much content appears to leave many students with reinforced misconceptions and a decreased ability to answer many astronomical questions" (p.54). Therefore, Sadler suggests that there are two main difficulties that learners encounter: their preconceptions and mastering easier concepts before learning the harder ones. Thus, Sadler emphasizes in his work that learners should undertake more difficult concepts only when mastery is achieved.

It is up to instructors to decide which topics to teach and which to omit, but there is one topic that cannot be excluded: stars. Stars and their properties are the most important part of any astronomy course (Prather, et al., 2009). Most information about stars can be determined from the Hertzsprung-Russell diagram (H-R diagram), a plot of stars specifying their temperatures and luminosities. In addition, the diagram not only displays relationships between absolute magnitude, luminosity, and temperature, but also shows the spectral classification of the stars. The sizes and masses of stars can be determined from the diagram as well. The H-R diagram is also used to represent stars evolutionary paths. With all this said, the H-R diagram is considered to be the most important diagram in stellar astronomy (Swinbank, 1997). However, it has been documented that students struggle with making sense of the diagram (Adams & Slater, 2000). Lightman and Sadler (1993) reported that approximately only 25% of high school students who attended astronomy courses could properly read the diagram.

Teaching properties of stars presents yet another set of challenges. The concept cannot be taught through direct experiments as most of other concepts in astronomy. For instance, students cannot put Saturn into a bathtub full of water to test the planet's density (Marschall, Snyder, & Cooper, 2000). And it would take many night hours, days, weeks, and even months just to observe the motion of Galilean moons before students could determine the mass of Jupiter. Learners also cannot readily observe pulsations of Cepheid variable stars from their backyards in order to find distances to other galaxies. And they most definitely cannot use professional multibillion dollar telescopes for six-months-long observations of stars in order to find their parallaxes and thus determine distances to the stars and thereafter plot them on the H-R diagram. This teaching and learning technique of hands-on observational astronomy, would lead to many years before the students' graduation date. City lights, weather conditions, time consuming observations, and expensive inaccessible equipment are common obstacles in teaching astronomy. Since astronomy is largely based on observations (Marschall, 2000) to which students cannot dedicate their entire time spent in school, simulated environments become indispensable tools for teaching and learning the subject.

II.II. Computer-Assisted Instructions in Science Classes

During the past couple of decades, computer-based technologies have started to transform both the teaching and learning of introductory science courses. As Sasha Barab (2003) points out, we live in a time where "pedagogical theory and technological advances have created an opportunity to design innovative and powerful environments to support learning" (p. 200).

In the past decade, a number of social sciences researchers have investigated the role of computer-based technologies in constructivist learning environments: Barab and Duffy (2000), Dede (2005), Edelson (2001), Gredler (2004), Hmelo (1998), Jonassen (1995, 2000), Pea (1985), Rieber (1990), and Shank (1999), among others. Across a wide range of studies investigators conclude that computer-based technologies can be used as cognitive tools to enhance the learning experience of students (Bransford, Brown, & Cocking, 2000; Kim & Hay, 2005) build motivation (Barab & Dede, 2007; Duffy & Cunningham, 1996; Nickerson, Corter, Esche, & Chassapis, 2007; Wieman, 2007) and curiosity (Cameron, 2003), augment student confidence (Adams, et al., 2006; Marschall, 2000), increase level of mastery of a topic and performance in a course (Gentry, Dickinson, Burns, McGinnis, & Park, 2006; Marschall, 2000), and amplify and reorganize student thinking processes (Jonassen, 1996; Hawkins & Pea, 1987; Pea, 1985). Simulations are a particularly powerful form of computer-based learning technology, and are commonly used for the teaching of more difficult and abstract concepts in science classes (Pea, 1985; Rieber, 2005; White & Frederickson, 1998).

In addition, the social aspect of working with interactive environments is hypothesized to be an important determinant of outcomes in educational visualizations, simulations, and tutorials. Collaborative work among peers is an opportunity for a dialog between the learners about the explored subject and task at hand. During social collaboration within peer groups, "students share the responsibility for thinking and doing" (Bransford, Brown, & Cocking, 2000, p. 184) as a result of these negotiations of the taken actions, learners display a high level of concentration on a topic that reflects on their own reasoning skills, problem solving, and performance on tests (Rao & DiCarlo, 2000).

Given the importance of stars as a topic in astronomy and the recognition of educational simulations as a powerful teaching tool in science a number of laboratory exercises have been developed. These laboratories allow learners to become familiar with various properties of stars, help students to measure the distances to stars and galaxies, find the sizes of celestial bodies, and roam the lands of distant planets (Marschall, 2000; Taasoobshirazi, 2006; Yair, 2003). Some of the most popular simulations currently used in college-level astronomy courses are: *Project CLEA*, *RedShift* College Edition, *Starry Night,* and *Stellarium*.

In order to aid students in their understanding of concepts in astronomy such as those related to the H-R diagram, computer educational simulations use is increasingly more frequent. For instance, in January 2008 *Project CLEA* introduced a new simulation – the *H-R Diagrams of Star Clusters*, to support students' ability to understand the diagram. The same year, in order to assist the students in their understanding of stars properties and the H-R diagram the Department of Physics and Astronomy at the University of Missouri developed the aforementioned *Stellar Properties Laboratory (SPLab)* project, a short description of which is given below.

The *SPLab* was designed to address a problem common to most laboratory projects and simulations - they are independent of one another and do not support students in their grasp of the connectivity among the various aspects of astronomy. The *SPLab* combines several basic concepts of physical properties of stars and requires students to connect these properties, by using one in order to find another one until learners' can identify types of the stars and plot them on the H-R diagram.

II.III. Stellar Properties Laboratory Project (SPLab)

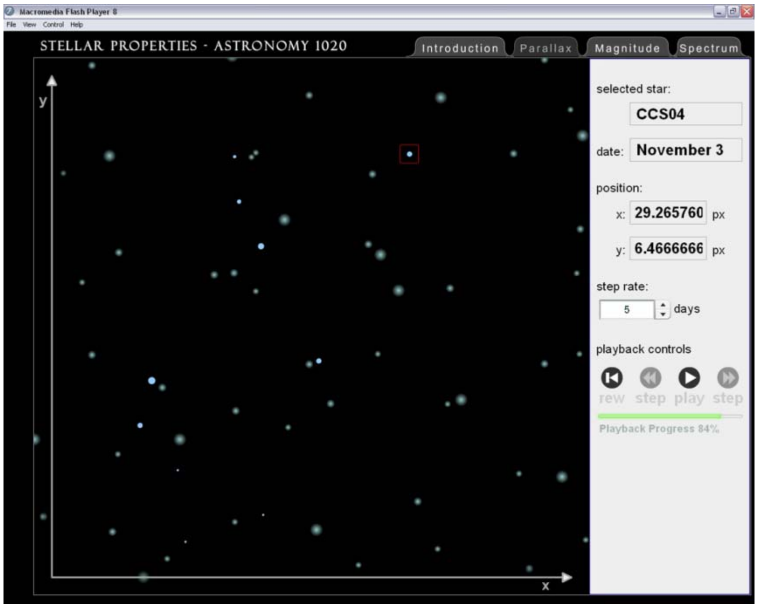
The *Stellar Properties Laboratory* project is designed for non-science major, college students. The goal of the project is to help students learn stellar properties (Ruzhitskaya & Speck, 2007). Not by doing separate exercises, but by becoming scientists through simulation and conducting a step-by-step investigation. Students are involved in collecting data through observations, analyzing the data, calculating stars' missing properties, and drawing conclusions. *SPlab* guides students along a path from finding stellar parallax to plotting stars on the H-R diagram. It helps students to piece together fragments of information from what they learned about the various properties of stars throughout the course.

During the work on the laboratory assignment, students are engaged in the following steps:

1. *Parallax and Distance.* The laboratory work starts with students' observations of a computer-simulated shift of seven star positions in the sky due to the Earths' revolution around the Sun. First, students must determine the time between observations, then, using the observed parallax angles, students can calculate distances to the given stars (Fig. 1.1).
2. *Apparent Magnitude, Absolute Magnitude, and Luminosity.* In this step students learn about photometry and collect photons emitted by the stars. Students take three simulated readings of photon counts and find the average to find the average flux of energy. Then they compare the found numerical value of the flux with a table of apparent visual magnitudes. Combination of the apparent magnitudes with the previously found distances allows students to calculate absolute visual magnitudes of the stars. As soon as students find absolute magnitudes, they can locate stars luminosities using a diagram provided to them.
3. *Spectrum, Spectral Class, and Surface Temperature.* The next observational property of the stars is stellar spectrum. In this step students learn to determine the stars spectral classes using spectra of the stars. From the spectral classes students determine stellar surface temperatures.
4. *Radius.* Using the luminosity and surface temperature values that were discovered earlier, the students calculate the stars radii.

5. *Star Type*. Having found temperatures, luminosities and radii of the sample stars, the students find what types of stars they have been observing. With this final piece of information, students can successfully plot the stars on the H-R diagram.

As discussed earlier in the chapter, connectivity among the concepts is important. Without seeing the whole problem, learners cannot assess importance and meaningfulness of a concept (Diakidoy, 2001). Thus, the simulated two-dimensional environment of the *SPLab* is designed to allow students to connect several astronomical concepts starting with a simple observation of stars moving across the sky – the stellar parallax, to the last step of plotting the stars on the H-R diagram. In order for students to be successful in this activity they must master the first simple concept of the stellar parallax.



selected star: CCS04

date: November 3

position:

x: 29.265760 px

y: 6.4666666 px

step rate: 5 days

Figure 1.1: This is a snapshot of a computer screen of *SPLab*'s first exercises - observing the stellar parallax and taking measurements of apparent shifts of stars.

However, an evaluation of students performance in the project conducted by the author in November-December 2008 showed that many students even with the support of the *SPLab* cannot successfully plot stars on the H-R diagram as they fail to complete the first step of the project correctly: calculate the distance to the stars, something which is achieved through the parallax method.

II.IV. Stellar Parallax Interactive Restricted and Unrestricted Tutorial (SPIRUT)

To help the students overcome difficulties in the *SPLab*'s first step a new guided introduction to the simulation has been developed. *SPIRUT*, is an interactive, inquiry-based tutorial that asks students to predict changes in the simulated environment as a result of the students' manipulation of the provided settings. For instance, the students are asked to predict how an apparent shift of a star in the sky changes if the size of the Earth's orbit is increased or decreased. The students are asked to reflect on what is observed (a more detailed description of the *SPIRUT* is included in Chapter 3, page 75).

It is important for students to understand the concept of the stellar parallax which is an apparent shift of stars as a result of the Earth's orbital motion. Parallax is found based on the Pythagorean Theorem of determining sides and angles of a right-angle triangle ($c^2 = a^2 + b^2$) where the sine (θ) of an angle is defined as the ratio between the opposite side and the hypotenuse: $sin(\theta) = opposite\ side\ /\ hypotenuse$.

The course textbook *Cosmic Perspectives* authored by Bennett, Donahue, Schneider and Voit, (2010) provides students a detailed description and explanation, graphical representation of finding the stellar parallax (see Fig. 1.2), mathematical relation between parallax and a distance (see Fig. 1.3), and a mathematical example of how to find a distance to a star using the stellar parallax.

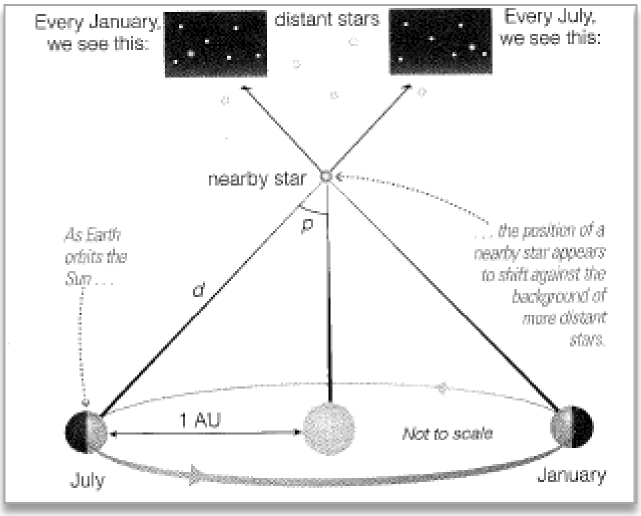
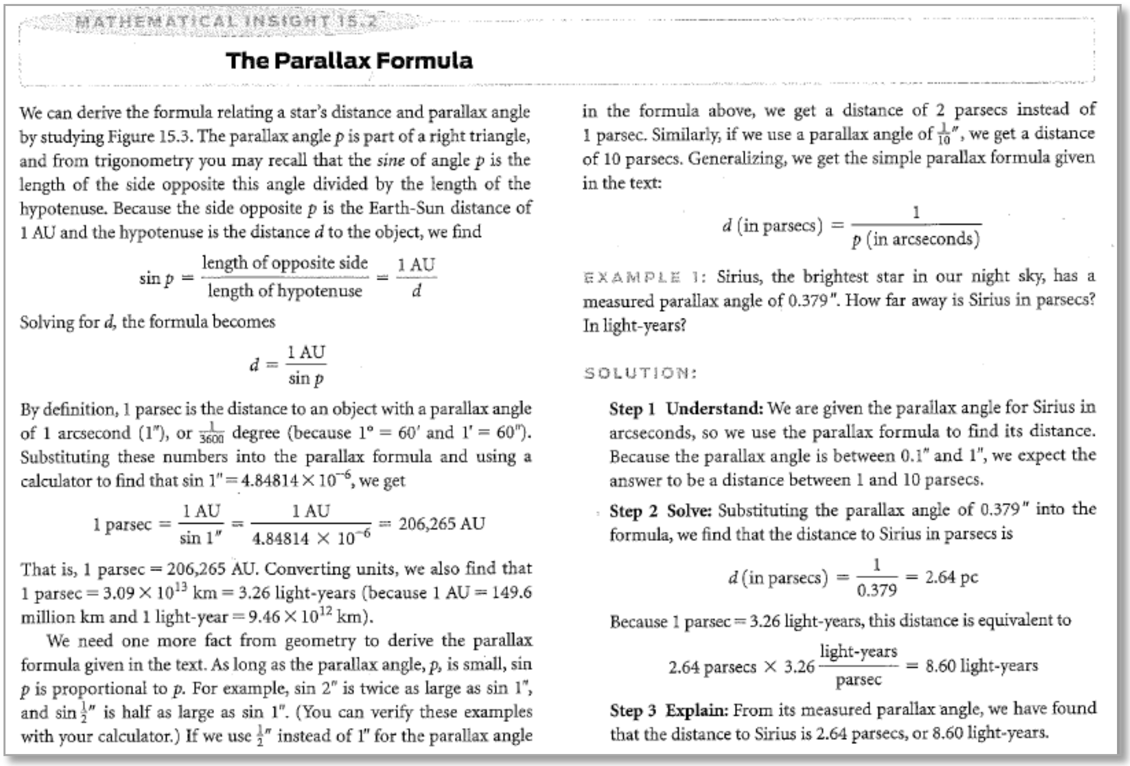


Figure 1.2: This figure from the course' textbook graphically represents the parallax concept. The notations under the figure explain: "Parallax makes the apparent position of a nearby star shift back and forth with respect to distant stars over the course of each year. The angle *p*, called the *parallax angle*, represents half the total parallax shift each year. If we measure *p* in arcseconds, the distance *d* to the star is *1/p*. The angle in the figure is greatly exaggerated: All stars have parallax angles of less than 1 arcsecond." (Bennett, Donahue, Donahue, Schneider & Voit, 2010, p.495)

Nevertheless, when students observe moving stars across the sky using the developed simulation they are not sure what dates of the observations they should choose for their calculations and how to use them in order to find the parallax. Students do not apply the material from the textbook to practical work on the project.

Following observations of how students work on *SPLab*, the course instructors suggested that an inquiry-based approach to the guide for the simulation may help students better explore the given situation and find the way to solve the problem. *SPIRUT* allows students not only to interact with the tutorial by using a number of available manipulatives, but also prompts them to draw conclusions based on answering posed questions, making predictions and reflecting on the observed changes as the result of manipulation of the variables. This scaffolding feature of the interactive tutorial is designed to help students grasp the concept and build their capacity for a mathematical application of the stellar parallax before they move on to work with the *SPLab* project (Fig. 1.4).



MATHEMATICAL INSIGHT 15.2

The Parallax Formula

We can derive the formula relating a star's distance and parallax angle by studying Figure 15.3. The parallax angle p is part of a right triangle, and from trigonometry you may recall that the *sine* of angle p is the length of the side opposite this angle divided by the length of the hypotenuse. Because the side opposite p is the Earth-Sun distance of 1 AU and the hypotenuse is the distance d to the object, we find

$$\sin p = \frac{\text{length of opposite side}}{\text{length of hypotenuse}} = \frac{1\ \text{AU}}{d}$$

Solving for d, the formula becomes

$$d = \frac{1\ \text{AU}}{\sin p}$$

By definition, 1 parsec is the distance to an object with a parallax angle of 1 arcsecond (1″), or $\frac{1}{3600}$ degree (because $1° = 60'$ and $1' = 60''$). Substituting these numbers into the parallax formula and using a calculator to find that $\sin 1'' = 4.84814 \times 10^{-6}$, we get

$$1\ \text{parsec} = \frac{1\ \text{AU}}{\sin 1''} = \frac{1\ \text{AU}}{4.84814 \times 10^{-6}} = 206{,}265\ \text{AU}$$

That is, 1 parsec = 206,265 AU. Converting units, we also find that 1 parsec $= 3.09 \times 10^{13}$ km = 3.26 light-years (because 1 AU = 149.6 million km and 1 light-year $= 9.46 \times 10^{12}$ km).

We need one more fact from geometry to derive the parallax formula given in the text. As long as the parallax angle, p, is small, $\sin p$ is proportional to p. For example, $\sin 2''$ is twice as large as $\sin 1''$, and $\sin \frac{1}{2}''$ is half as large as $\sin 1''$. (You can verify these examples with your calculator.) If we use $\frac{1}{2}''$ instead of 1″ for the parallax angle in the formula above, we get a distance of 2 parsecs instead of 1 parsec. Similarly, if we use a parallax angle of $\frac{1}{10}''$, we get a distance of 10 parsecs. Generalizing, we get the simple parallax formula given in the text:

$$d\ (\text{in parsecs}) = \frac{1}{p\ (\text{in arcseconds})}$$

EXAMPLE 1: Sirius, the brightest star in our night sky, has a measured parallax angle of 0.379″. How far away is Sirius in parsecs? In light-years?

SOLUTION:

Step 1 Understand: We are given the parallax angle for Sirius in arcseconds, so we use the parallax formula to find its distance. Because the parallax angle is between 0.1″ and 1″, we expect the answer to be a distance between 1 and 10 parsecs.

Step 2 Solve: Substituting the parallax angle of 0.379″ into the formula, we find that the distance to Sirius in parsecs is

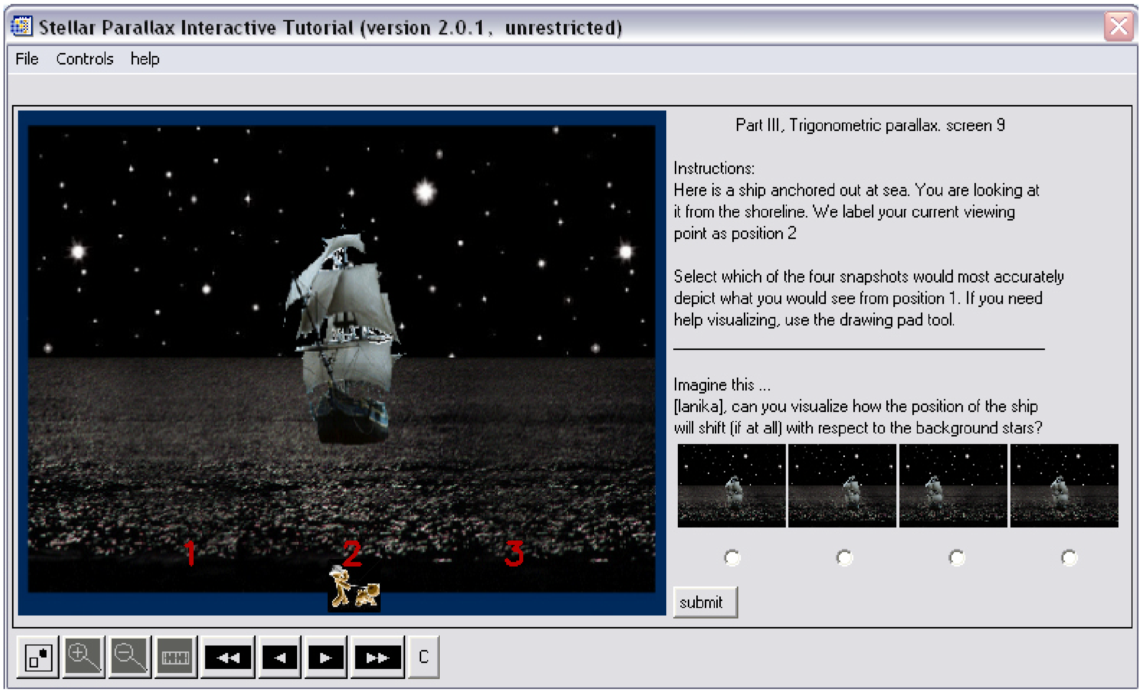
$$d\ (\text{in parsecs}) = \frac{1}{0.379} = 2.64\ \text{pc}$$

Because 1 parsec = 3.26 light-years, this distance is equivalent to

$$2.64\ \text{parsecs} \times 3.26\,\frac{\text{light-years}}{\text{parsec}} = 8.60\ \text{light-years}$$

Step 3 Explain: From its measured parallax angle, we have found that the distance to Sirius is 2.64 parsecs, or 8.60 light-years.

Figure 1.3: The figure shows mathematically represented relation between distance to a star and stellar parallax. (Bennett, Donahue, Donahue, Schneider & Voit, 2010, p.495)



Part III, Trigonometric parallax, screen 9

Instructions:
Here is a ship anchored out at sea. You are looking at it from the shoreline. We label your current viewing point as position 2

Select which of the four snapshots would most accurately depict what you would see from position 1. If you need help visualizing, use the drawing pad tool.

Imagine this ...
[Ianika], can you visualize how the position of the ship will shift (if at all) with respect to the background stars?

○ ○ ○ ○

Figure 1.4: This screenshot of an interactive exercise of *SPIRUT*. Students have to predict an apparent shift of the ship as it is viewed from different positions on the seashore.

To better explore the effect of an inquiry-based guide to learning, the restricted, program-controlled version of *SPIRUT* was created. The non-inquiry-based guide version of the simulation provides the same manipulatives and visual representations but it lacks the scripted prompts for prediction and reflection of the inquiry-based scaffolding which shapes students' actions in the inquiry-based guided simulation. The guided version of *SPIRUT*, is a step-by-step scaffolding tool that restricts students' control of their actions to a set of activities asking the students questions, guiding their conduct of simulated experiments, observations, and soliciting reflection on the actions taken.

III. Rationale for the Study

As it has been demonstrated, there are a number of challenges that learners face in astronomy courses: large number of concepts to be learned in a relatively short period of time, more difficult concepts are introduced even though the basic ones may not yet be mastered, not having deep prior knowledge and even possessing misconceptions, and lack of active learning in a form of direct experiments for enhancing learning experience. Therefore, even though students attend lectures, read textbooks, and watch animated visualizations during the classes, they still struggle to piece information together and thus their knowledge remains disconnected.

The current study hypothesizes that with an appropriate tool such as a computer-supported tutorial with include realistic-looking scenes for creating an authentic experience (Gredler, 2004), and scaffolding for helping students to master the concept in a step-by-step inquiry-based process (Hmelo & Guzdial, 1996) along with negotiation of meaning during peer interaction in an active-learning environment (Rao & DiCarlo, 2000), building deeper conceptual understanding of the subject and abstract thinking skills can be achieved.

III.I. Research Questions

The investigation aims to answer the following research questions:

1.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

  1.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

  1.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

2.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions in student-controlled or tutorial-controlled *SPIRUT*?

  2.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions in student-controlled or tutorial-controlled *SPIRUT*?

  2.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions in student-controlled or tutorial-controlled environments of *SPIRUT*?

3.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are supported by working with a partner or working independently?

  3.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are supported by working with a partner or working independently?

  3.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are supported by working with a partner or working independently?

III.I.I. Hypotheses to be Tested

1.1. Students who were involved in active learning through their direct interactions with the computer-based tutorial by means of a set of tools and were exposed to visual static and animated representations of the concept, would construct greater knowledge of the stellar parallax concept.

  1.2. Students who were involved in active learning through their direct interactions with the computer-based tutorial by means of a set of tools and were exposed to visual static and animated representations of the mathematical representation of the concept, would perform simple mathematical computations and remember facts underlying the stellar parallax concept better than students who were performing similar calculations in the paper-based tutorial.

  1.3. Students who were involved in active learning through their direct interactions with the computer-based tutorial by means of a set of tools and were exposed to visual static and animated representations of the concept, would learn how to reason their answers and transfer knowledge better than students who worked with the paper-based tutorial.

2.1. Students who worked with program-controlled *SPIRUT* would construct greater knowledge of the stellar parallax concept as a result of *SPIRUT*'s tailored path for the students based on their actions.

    - 2.2. Students who worked with program-controlled *SPIRUT* would learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better as a result of *SPIRUT* tailored path for the students based on their actions.
    - 2.3. Students who worked with program-controlled *SPIRUT* would learn how to reason their answers and transfer knowledge better as a result of *SPIRUT* tailored path for the students based on their actions.

- 3.1. Students who worked with a partner would construct greater knowledge of the stellar parallax concept as a result of their cooperative work with the peer.
    - 3.2. Students who worked with a partner would perform simple mathematical computations and remember facts underlying the stellar parallax concept better as a result of their cooperative work with the peer.
    - 3.3. Students who worked with a partner would learn how to reason their answers and transfer knowledge better as a result of their cooperative work with the peer.

IV. Professional Significance of the Study

In recent years, calls for improved outcomes from science education and moves to more constructivist approaches to education have motivated a number of investigations on how students construct understanding in astronomy courses (Bailey, 2004). Additionally, rapid advances in technology bring an opportunity to harness the new technology capabilities and their accessibility to the service of science education. The proposed study addresses the need and opportunity to research methods for improving the role of technology in astronomy education.

The study investigates the role of inquiry-based learning methods on students' construction of conceptual knowledge of scientific concepts and their development of abstract thinking skills when learning from computer-assisted instruction in science courses. The study also explores how students' learning is impacted through peer interaction during the use of a tutorial. Thus, the study aspires to contribute new knowledge to the fields of science education and technological innovations for science education.

V. Key Definitions

Abstract Thinking. The thinking is characterized by one's aptitude to conceptualize: the ability to generalize from a concrete experience as well as ability to apply an abstract concept to new situations and/or surroundings including mathematical applications (Clements & Sarama, 2004). There are many ways in which abstract thinking skills are applied to transforming real world situations in mathematical representations. For instance, Johannes Kepler, using recorded observations of apparent retrograde motion of planets, determined and mathematically proved that such motion must be produced not by perfect circular but by elliptical orbits (Hirshfeld, 2001). The ability of children to recognize geometrical patterns as well as their ability to divide slices of an orange among friends or a number of sweets among different groups of individuals are all examples of the skills development and application of abstract thinking (Bryant & Squire, 2001).

Inquiry-based Learning. Also known as inquiry-based science is a theoretical framework for a constructivist approach to active learning where students build their knowledge by means of guided or unguided inquiries and negotiate the construct meaning with peers (Colburn, 2000; Duffy, Dueber, & Hawley, 1998).

Manipulatives. In a computer-based interactive environment, a set of tools that allows a user to perform desired actions resulting in immediate change in the environment and as seen on a computer screen (Gredler, 2004; de Jong and van Joolingen, 1998).

Peer Interaction. A peer learning approach in which learners construct knowledge while working on a given task collaboratively exchanging ideas, sharing prior knowledge, opinions, and making decisions (Hooper, 1992; King, 2002).

Scaffolding. In a constructivist learning environments scaffolding is a method of supporting students' performances. The scaffolding aids students' building of conceptual understanding based on their prior knowledge and fades out when support is no longer needed (Hmelo-Silver, 2006; van Joolingen, de Jong, & Dimitrakopoulou, 2007).

SPIRUT – Stellar Parallax Interactive Restricted and unrestricted Tutorial. The science inquiry computer-based adaptive interactive tutorial created as an introductory portion of the *Stellar Properties Laboratory* project (*SPLab*) or a stand-alone tutorial to scaffold students learning of stellar parallax. *SPIRUT* includes a set of manipulatives for simulated observations and data collection, change of variables, question prompts in a form of multiple-choice and essay questions, and text fields for writing reflections on the observations. *SPIRUT* integrates two modes: learner-control and program-control options (see Chapter 3 for more details).

Stellar parallax. An apparent shift of nearby stars and a method used in astronomy to find distances to these nearby stars by observing their apparent shift (parallax) on the background of more distant stars from two different points of the Earth's orbit six month apart. The relationship between the distance to a star, baseline of the observations and an apparent shift of the star can be represented by the following equation $d = 1/p$, where d is the distance to the star (measured in parsecs), 1 is the distance between the Sun and the Earth (measured in astronomical units) and p is a parallax angle (measure in arcseconds) and corresponds to the apparent shift is ½ of the shift.

Transfer of knowledge. A process in which learners can apply their knowledge in situations different from their learning environment in which this knowledge was originally constructed (Pellegrino, Chudowsky, Glaser, 2001).

VI. Limitations of Study

To assess students' conceptual gain and retention of their constructed knowledge of the stellar parallax concept, three assessment tests using *SPA* were administered within an eight-week period. A pre-test and post-test were given students within a two-day period and a retention test was administered eight weeks later. To compare students' performance on the assessment tests the same set of questions was used. Even though, the questions were presented in different orders on all three tests, it is possible that students instead of answering a question based on their current knowledge of the concept were basing their answers on recollection of their previous answers.

Another limitation of the study is a threat of the internal validity of the study is experimenter expectancy or communication to the participants of expectations of the study (Neuman, 1997). Since the study investigator is a developer of the tutorial, it is possible that the investigator indirectly communicates the hypotheses of the study through her behavior, or leading questions during follow-up interviews, and/or observations. The likelihood of this occurrence is minimized by implementing semi-structured interviews with a set of the written questions for the qualitative data collection during the follow-up interviews (Appendix D) and the use of three co-raters for assigning scores on each assessment test and another co-rater for an analysis of conducted interviews.

Since the study was conducted using a group of non-science major college students based on the tutorial designed as a part of their introductory astronomy course, the generalization of the study must be approached with caution. The replication of a study among other groups of people may produce similar or different results based on the age of the participants, their level of mathematical abilities, prior knowledge, and learning goals, to name but a few contributing factors.

CHAPTER 2

LITERATURE REVIEW

I. Introduction

This chapter reviews literature on inquiry-based learning with an emphasis on implementation of the theory in science classes and its use as a framework for designing supporting tools for computer-based educational tutorials. It also describes the emerging *Technology-Supported Inquiry Learning* theory (*TSIL*) and its role in creating apt computer-based educational tools. The chapter emphasizes the role of mental dual coding of information and its affect on human memory as it is described by *Dual Coding Theory* (*DCT*). The chapter discusses previous studies and their findings on implementation of the interactive tools embedded in various types of computer-assisted instruction, with a goal of aiding students' cognitive processes and enhancing their learning experience. The presented studies illustrate a number of beneficial aspects of using various scaffolding tools, social interactions among peers, and argues for the importance of using specifically designed tools for particular learning goals to be achieved by educational computer-based interactive environments.

II. Inquiry Learning: Background

The theoretical framework of the study rests on the *TSIL* theory (Edelson, Pea, & Gomez, 1996) that advocates use of particular technological tools for particular tasks based on the tasks learning goals. The thesis of the proposed research is that learners will better understand the stellar parallax concept and translate it into mathematical application/formula when the learners are scaffolded by appropriate tools and are involved in peer interaction.

The *TSIL* theory is rooted in the active learning approach for knowledge construction advocated by John Dewey and Lev Vygotsky in the first half of the 20th century. The work of both social science researchers has led to an appreciation that learners should be active participants within collaborative settings (Bransford, Brown, & Cocking, 2000; Dewey, 1916; Duffy, Dueber, & Hawley, 1998; Jonassen & Land, 2000; King, 1998; Orrill, 2001; Romiszhowski & Mason, 2001; Solomon, 1983; Vygotsky, 1978).

In addition, John Dewey (1938) argued for the importance of authentic settings for learning. He suggested that learning takes place when a person is taken out of the comfort of a habit and put in a situation where a decision must be made through the process of reasoning. In addition to experiential learning (Dewey, 1916), the philosopher believed in the importance of social interconnection. He pointed out that “isolation of subject matter from a social context is the chief obstruction in current practice to securing a general training of mind” (Dewey, 1916, p.79). By “general” Dewey meant the “broad and flexible” (Dewey, 1916, p.78) mindset. Lev Vygotsky also argued the importance of social collaboration in educational settings (Vygotsky, 1997) in support of active learning. Vygotsky believed that an individual is a social being who constructs his/her knowledge based on the historical culture of a social group to which he/she belongs through mediation of language and symbolic representations within the group. The social interaction during participatory learning process was explained by Vygotsky (1978) in his concept of the zone of proximal development (*ZPD*). The social scientist explained the *ZPD* as, "the distance between the actual developmental level as determined through independent problem solving and the level of potential development as determined through problem solving under adult guidance or in collaboration with more capable peers" (p. 86). This mediated learning provides learners with needed support to solve simpler problems before continuing to more difficult ones. Thus, the psychologist believed that assisted collaborative active learning allows learners to reach their maximum level of development. Vygotsky’s idea of bringing these social mediation activities to a classroom found many supporters in the American educational community (Glassman, 2001).

Dewey (1938) and Vygotsky (1997) inspired Glassman's (2001) argument that interest drives inquiry and must come from interaction between a learner and the present situation. Thus, interest motivates a learner to make sense out of presented phenomena in order to find a solution to an arisen problem (Dewey, 1938). According to Dewey (1938), during this process of making a decision reflection on the situation and reasoning take place.

Following from the early work of Dewey and Vygotsky, the contemporary constructivist approach to a learner-centered environment, calls for learning activities in which participants experience situated authentic contexts. Students are also introduced to multiple perspectives that empower learners to negotiate and to define a meaning based on their personal beliefs and previous knowledge (Bransford, et al. 2000; Dewey, 1916). In this environment students are learning by doing (Dewey, 1938; Shank, 1999); through active participation in the learning activity students are engaged in collaboration and in the processes of inquiry, discovery, and innovation. As a result, active learning is a constructivist approach to teaching and learning that allows students to take control over their own learning process (Bell & Kozlowski, 2007; Bransford, et al. 2000). At the same time, learning can be "directed toward understanding through exploration and experimentation" (Bell & Kozlowski, et al., p.299) where the inquiry process is based on deductive (Colburn, 2000) and inductive (Dewey, 1938) reasoning.

However, it is interesting to notice, that Kirschner, Sweller, and Clark, (2006) while referring to discovery and inquiry learning, suggested that unguided and experientially-based approaches do not necessarily promote learning. In response, other researchers pointed out that there are several approaches to inquiry such as structured, guided, open, and learning cycle (Colburn, 2000; Hmelo-Silver, Duncan, & Chinn, 2006; Ross & Morrison, 2001) that are generally used in inquiry learning. In guided inquiry, students are provided with materials to investigate a problem where it is up to them to find a right solution (Colburn, 2000), make plans and procedures for arriving at a solution and take control over the learning process. This type of inquiry learning originates in the practices of scientific inquiry where the processes of questioning, collecting and analyzing data, drawing conclusions, and developing arguments are benchmarks of the learning processes (Kuhn, Black, Keselman, & Kaplan, 2000; Krajcik & Blumenfeld, 2006). Inquiry learning engages students in cognitive processes of sense making that requires developing explanations based on observed evidence, reflecting and communicating ideas (Barab & Duffy, 2000; Fraser, 1998; Hmelo-Silver, 2006; van Joolingen, de Jong, & Dimitrakopoulou, 2007).

The requirement that inquiry learning makes sense out of presented information finds wide implementation in science classes. Today, researchers have a body of evidence supporting the thought that a simple explanation of a phenomenon does not necessary help students to understand the phenomenon (Bransford, et al. 2000). Vosniadou and Brewer (1992) found in their study, conducted with 60 middle school children, that a science teacher's explanation of why Earth is a sphere, did not help students to understand why this was indeed the case.

To help students better understand scientific concepts, technology has been implemented in classrooms as a learning tool in a variety of ways. For more than 20 years computer-based interactive environments and scientific visualizations have been used with the goal of creating engaging environments for students of science (van Joolingen, et al. 2007).

II.I. Dual Coding Theory (*DCT*)

The dual coding theory is based on verbal and non-verbal processing of information. The theory was developed by Alan Paivio in 1986 who suggested that since human cognition is simultaneously processing language, images and events by coding them accordingly: "The language system is peculiar in that it deals directly with linguistic input and output (in the form of speech or writing) while at the same time serving a symbolic function with respect to nonverbal objects, events, and behaviors. Any representational theory must accommodate this dual functionality" (Paivio, 1986, p. 53). People construct knowledge through the process of recognition and analysis of perceived information as a result of an event segregated into words and images – verbal (words) and visual (images) constructs of the coding process.

The theory embraces this idea and states that people encode and later retrieve processed information using verbal and imagery systems consisting of words and images (logogens and imagens). These two cognitive subsystems can be activated and work on different levels:

- Representational: One can be active without another
- Associative: Both systems can be activated in parallel
- Referential: Information flows from one to another

Any given word has its visual representation in our cognition in a visualized physical form or its written representation. When a word is heard for the first time, cognition creates an associative chain of used sounds, image, and written visual representation with it (if the word was demonstrated in writing). For instance, when a desktop computer is shown and a word “computer” pronounced, it could be coded as a sound, written letters in their particular order and an image of a computer on a student’s desk. When a familiar word is pronounced, the human mind immediately draws upon prior knowledge and associations while retrieving the word from memory. Thus, memory is a crucial component of the theory “because it is the basis of all knowledge and thought” according to Allan Paivio (2006, p. 4).

Logogens and imagens also have different structures. As shown by a Stella' modeled visual representation of the theory (Fig. 2.1), the verbal subsystem (logogens) is more of a hierarchical structure where one constituent of a logogen depends on another. The nonverbal (imagens) subsystem is much less restrained and any image can be recalled independently from the whole picture (or set of imagens that creates the picture in memory). Such a structure of interconnections between the two subsystems has a dual character as well. It can facilitate building a level of understanding and processing of information as well as create obstacles when intense processing of imagens begins to interfere with logogens.

What type and how many system connections will be triggered depends on the stimuli. The theory identifies five sense modalities: visual, auditory, haptic (a sense of feel), gustatory (a sense of taste), and olfactory (a sense of smell). While imagens can relate to all five of them, logogens are affiliated only with the first three (Tab. 2.1).

Table 2.1: *This relationship between specific modalities and mental codes is adapted from Paivio (Paivio 1986) and Sadoski (Sadoski 2001).*

Symbolic Systems (Mental Codes)		
Sense Modality	**Verbal**	**Nonverbal**
Visual	Visual language, words (writing)	Visual objects
Auditory	Auditory language, words (speech)	Environmental Sounds
Haptic	Braille, handwriting	“Feel” of objects
Gustatory	-	Taste memories
Olfactory	-	Smell memories

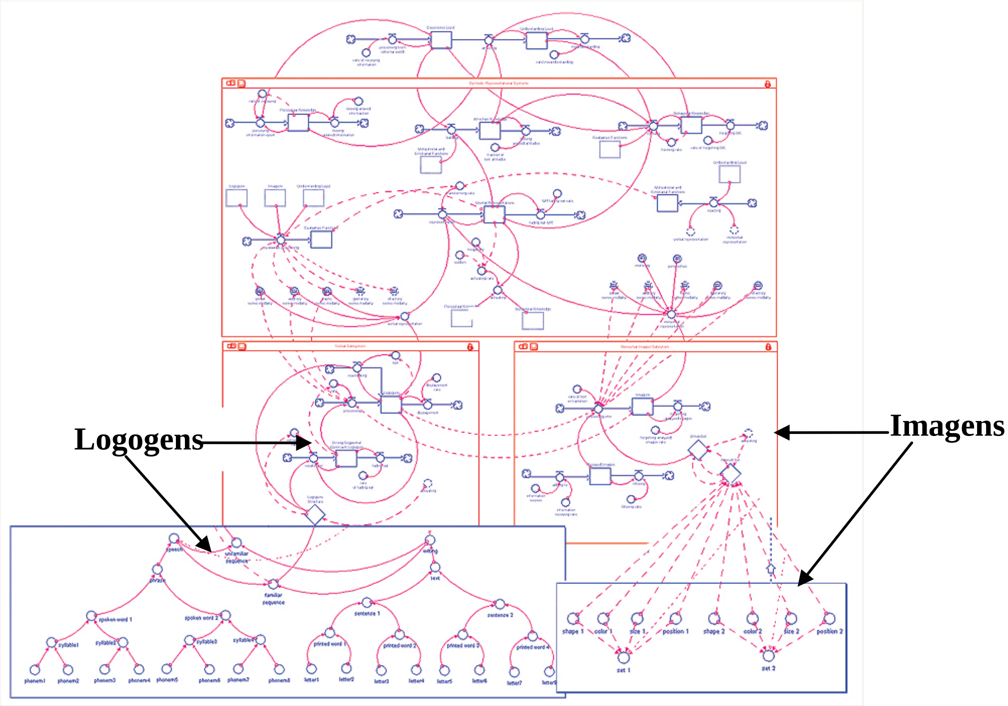


Figure 2.1: STELLA model of the Dual Coding Theory (*DCT*) shows the process of mental coding, storing and retrieving of information by a learner.

The visual dual coding theory model shows collaborating sections of symbolic representational systems and verbal and nonverbal (imagen) subsystems. The verbal subsystem of the model reflects the hierarchical structure of logogens and the formation of strong sequential constraint logogens that are formed due to frequent use of units of speech and text. The nonverbal cognitive subsystem does not have such strong constraints and has different relationships between perceived and analyzed images. The activated units of images can be retrieved independently.

The two displayed functions are involved in processing information built and distributed by imagery and verbal processes. Evaluative functions depend on logogens, analyzed images, and current level of understanding, and are influenced by all five modalities. The motivational and emotional functions, on the other hand, depend on representations of verbal and nonverbal subsystems, and are related to the current level of understanding.

Originally *DCT* was developed to explain importance of imagery in learners' perception of words. Therefore, the theory was more frequently applied in linguistics. However, today the dual coding theory is often applied in educational psychology for understanding how learners process information, construct their conceptual knowledge, build spatial skills, and work with problem-solving tasks (Paivio, 2006).

Mayer and Sims (1994), for example, applied the dual coding theory in their investigation of how to help students use verbal and visual information to understand scientific concepts. In their study the researchers compared two groups of students with high- and low- spatial abilities on perceiving concurrent and successive multimedia information. The results of the study showed that students with high-spatial abilities performed 50% better on transfer problems when subjected to concurrent multimedia instruction rather than successive instruction. At the same time, low-spatial ability students' performance had nearly no difference between the two forms of instruction. In addition, Mayer and Sims (1994) suggest that inexperienced students would receive the greatest benefits from concurrent multimedia instruction, as students with prior domain knowledge can draw on their previously coded mental representations as they read or listen and thus do not require further visual stimuli.

Another study was conducted by Suh and Moyer-Packenham (2007). The researchers examined application of dual coding theory in a number of multi-representational virtual mathematics environments where they investigated how young students build their algorithmic thinking. Thirty-six third-graders were divided into two groups: one group of students learned fractions using physical manipulatives and learned algebra using virtual manipulatives, while another group of students used virtual manipulatives to learn fractions and physical manipulaties for their algebra lesson. The investigators found that students from both groups who received virtual manipulatives treatment for both fractions and algebra lessons performed better on assessment tests. Based on the findings, the researchers concluded that “dual-coded representations in virtual manipulative environments, that combine visual images with symbolic notation systems, have the potential to be effective in teaching mathematical processes” (Suh & Moyer-Packenham, 2007, p. 4-215). They also suggested that as a result of dual coding of visual and verbal representations students will be able to learn even more complex algorithmic applications.

II.II. Technology-Supported Inquiry Learning (*TSIL*)

Although, the early use of learning technologies focused on transmitting the teacher’s instructions in a more engaging way to a larger number of students (Edelson, Pea, & Gomez, 1995; Wilson & Redish, 1989), today technological advancements in computers not only allows placing computers in classrooms and college laboratories (Edelson, Douglas, & Pea, 1999; Edelson, 2000) but also lets students use personal laptops for simultaneous wireless access to information while listening to lectures in large lecture halls and participating in explorations in computer-generated online learning environments. This advancement creates new opportunities to enhance students’ learning experience by making them active participants instead of passive listeners. Engaging students in a process of finding answers to the questions posed by an instructor, completing an experiment related to the theory that has just been explained – are just a couple of examples of how in-class interaction supported by computer-based technologies can enhance students' experience and support their knowledge construction (Edelson, et al. 1995; Lakkala, Lallimo, & Hakkarainenvan, 2005; Manlove, Lazonder, & Jong, 2006; Orrill, 2001; Salovaara, 2005; van Joolingen, et al. 2007).

Blumenfeld et al. (1991) identified six main aspects of technologies for supporting students' engagement in learning:

- enhancing learners' interest
- providing learners' access to information
- active and manipulable representation
- structuring the working process
- diagnosing and correcting errors
- managing complexity and aiding production.

A few years later, Edelson, Gordin, and Pea (1999) identified five main challenges to the successful implementation of inquiry-based learning that correspond to the aspects identified by Blumenfeld et al. related to learning technologies. The researchers highlighted the following challenges:

- learners' motivation
- accessibility of investigation techniques
- learners' background knowledge
- management of extended activities
- the practical constraints of the learning context.

Furthermore, Edelson et al. (1999) point out that “failure to address any of these challenges successfully” can not only prevent students from a needed level of engagement for meaningful investigation but may even “undermine learning” (Edelson, Gordin, & Pea, 1999, p.399).

Edelson and colleagues suggested that technology could be used to address and help to solve the identified challenges of supporting inquiry-based learning. As a result of several years of work, they proposed a theoretical formulation labeled the Technology-Supported Inquiry Learning theory (*TSIL*) which advocates integration of technology and activity design. The theory is based on the idea of creating particular tools for particular contexts in consideration with social interactions. The theory states that tools specifically designed for particular environment and social activity should support learners’ performance, which in turn will impact the learners’ knowledge gain (Edelson, Gordin, and Pea, 1999; Lakkala, Lallimo, & Hakkarainenvan, 2005; van Joolingen, et al. 2007).

Therefore, the technology-supported inquiry learning theory is not based on use of technology for mere transmission of ideas or human-computer interactions, but based on the nature of these interactions. The theory can be applied as a framework for creating technological cognitive tools to enhance learners’ performance through augmenting levels of their computer and social interactions. This facilitates the knowledge building process (van Joolingen, et al. 2007) by means of providing scaffolds such as investigation tools, knowledge resources, and record-keeping tools (Edelson, Gordin, & Pea, 1999). The theory has three objectives: help learners to develop (1) general inquiry abilities, (2) inquisition of specific investigation skills, and (3) understanding of science concepts and principles (Edelson, Gordin, & Pea, 1999).

Given this theoretical base, Ton de Jong suggested a set of learning inquiry processes that provide “a suitable basis for describing scaffolds specific to the learning process” (de Jong, 2006, p.108):

- orientation
- hypothesis generation
- experimentation (experiment design, prediction, data interpretation)
- drawing a conclusion
- making an evaluation.

Each of the named processes is particular to specific learning objectives and requires a particular cognitive tool(s) to support the efforts of the learners. These tools can be often found integrated in educational simulations, tutorials and other forms of interactive environments that are designed to support the users’ learning which takes place when learners are engaged in making sense out of presented phenomena – build meaning, construct knowledge and understanding of the subject (Barab & Duffy, 2000; van Joolingen, et al. 2007).

Hence, it can be concluded, that successful computer-based interactive environments developed for inquiry-based learning courses must include tools that adequately support students’ engagement while maintaining their curiosity, enable students to conduct investigations by providing necessary scaffolds to help students achieve learning goals, and encourage learners’ social collaboration.

III. Computer-based Interactive Environments in Science Classes

Constructivist learning environments allow students to not only learn by doing (Shank, 2002) but gain, fundamental techniques and skills such as collecting data and undertaking the analysis process which are used in modern astronomy (Marschall, 2000). At the same time, simulated environments in astronomy courses enabled by modern computer technologies not only provide access to virtual telescopes and deliver needed data for learning at anytime, regardless of weather conditions and time of day or night, but allow students to immerse into the experience of being a professional astronomer (Marschall, 2000). These “experiential exercises” (Gredler, 2004, p.571) have the ability to transport learners to another - simulated reality either it is a high-fidelity simulation or an educational tutorial.

This type of authentic learning requires of computer-based technologies to be able of creating a believable learning environment. These types of learning environments are especially beneficial to students in science classes where learners cannot participate in direct experiments. Astronomers and astrophysicists do not wait for 26,000 years to observe how stars will change in the sky as a result of the Earth’s precession and they do not go back in time to find out how was the Moon formed. Scientists do not travel to cores of the stars to learn about nature of their energy production neither they physically explore stellar remnants to learn their chemical composition, temperature, pressure, mass and many other aspects.

Most of the science today is done on computers. Computer-based learning environments create models that virtually recreate birth and deaths of stars, formation of planets, galactic collisions, and expansion of the universe. All scientists need is proper computer tools. So do students. As Jonassen (1995) states, “The tools do not control the specialist. Neither should computers control learning. Rather, computers should be used as tools that help learners to build knowledge” (p.44). Using properly designed computer tools: educational simulations and tutorials, students can become scientists and perform their own experiments and observations within a course’s allotted time. A growing tendency of teaching astronomy inside the classroom using computer screens instead of ground-based and space telescopes does not come as a surprise.

III.I. Visualization and Interactivity

One of the important aspects for maintaining learners’ interest in computer-supported activities is visual representations. The role of images and various forms of visual representations in educational settings has been studied for many decades and is now a well-documented area of research; especially in the fields of psychology, science education, and computer science.

Lloyd Rieber (1990) conducted a study with elementary school children to investigate the effectiveness of animated graphics to demonstrate Newton’s first law of motion. The study compared three groups of children: one group used animated simulation, a second group used static graphics, and the third, a control group, had no visual representation of the lesson at all. The researcher found that the animated graphics group was superior to static graphics and to the no graphics groups (Rieber, 1990).

Taking the role of graphical representations research further, Mayer, Mautone and Prothero (2002) set a goal to find out how to design a simulation for maximum development of thinking and required skills for geology-major college students. The investigators found that, when used in interactive simulations, visual representations were significantly superior to verbal prompts for the students' development of conceptual understanding and reasoning skills. However, taking into account the particularity of the context of the tasks (such as learning to identify certain types of rocks), the superiority of animated graphics probably depends on the learning objective. For example, the results of the study may not be the same if it were conducted in an English literature course.

Parker and Haywood (1998) examined the value of visual models and asked 63 education major students and first-year service teachers to explain simple concepts of changes of night and day, the seasons, and phases of the moon. Pre-test results showed a lot of confusion among the future teachers. After giving the students a treatment that included visual models, all students demonstrated significant change in their conceptual understanding of the events. The researchers came to the conclusion that there is a need for visualization in astronomy classes. They also pointed out the importance of interaction with the systems. The researchers concluded that for a specific subject, such as astronomy, the spatial 3D representations are preferable to the conventional 2D graphics.

Another computer-supported interactive learning environment was designed by the Center of Education Technology in Tel Aviv (Yair, Schur, & Mintz, 2003). A *Thinking Journey to Mars* was based on a visual representation of Mars and included simulated travel in space to the planet, fly-by, and surface terrestrial features. In their theoretical paper the authors of the project argued that high quality images with combination of 3D representation and interactivity based on assigned tasks would increase students' efficacy in the experience (Yair, et al., 2003) and will help learners to construct a scientifically accurate representation of the Martian world.

Chen, Yang, Shen, and Jeng (2007) investigated the influence of an interactive environment on learning astronomy topics in a guided discovery setting. The study was conducted among 21 six-grade Taiwanese students using *Desktop Virtual Reality Earth Motion System* (*DVREMS*). The learning environment was designed for demonstration of the Earth's orbital motion, rotation, and the Earth axis' tilt. *DVREMS* included a minimum set of functions and manipulatives. However, it allowed students to view the planet from different angles, change viewpoints, adjust the Earth's tilt, and stop its rotation. The researchers received mixed results. While students scored overall higher on the post-test, they did struggle with the concept of day and night change as a result of the Earth's rotation. The researchers attributed this shortcoming to *DVREMS* as not being a fully immersive environment.

Computer-based interactive environments can be very involving and engrossing - an important factor in engaging students in performing a task. Nevertheless, while full engagement is good for motivation, experience, and goal-achieving progression, it can also hinder the educational outcome (Lim, 2006). Learners completely engulfed in interactions may not have time to reflect on what they are doing and why (Barab & Duffy, 2000; Bodemer, Ploetzner, Feuerlein, & Spada, 2004).

For example, simulations designed using popular Adobe Flash software, as pointed out by Adams and her collaborators (2008) are fun and engage students in explorations, but at the same time, some features can be too much fun to play with and as a consequence distract students from learning. Researchers came to this conclusion while they examined 61 Physics simulations.

In addition, in order to maintain students' engagement in interactivity and keeping their concentration on a learning task simultaneously, an alternative approach of explicit constraining of users' interactions was suggested by social science researchers (Adam et al., 2008; van Joolingen & de Jong, 1998). As of today, this approach is not yet sufficiently investigated and supported by empirical studies. However, this approach relates to the current investigation since it was based on a tutorial with a program-control option which constrained learners' freewill actions through the set of consequent steps requiring students' predictions and reflections in order to advance further.

While there are a number of studies suggesting that interactivity and visual representations are important components of successful educational computer-based environments, careful balance of these components is important for achieving the educational value of these programs.

III.II. Complexity and Scaffolding

Another aspect that draws researchers' attention is the level of an interactive environment's complexity: an exceedingly complex learning environment requires extra effort and time to complete a presented exercise, which may result in diminished learning (Bodemer et al., 2004; Edelson, Pea, & Gomez, 1995; Lim, 2006; White, 1984). Bodemer et al. (2004), in his study with 81 students, showed that complicated computer-based environments demand a lot of mental processing, learners start paying attention to the environment's salient features that do not always contain educational value. Users stop interacting in a structured goal-oriented way and, as a result, do not piece together provided information. Learners may become overwhelmed by information and be pressed by time to progress without analyzing the sequence of events or reflecting on their actions. Hence, learners become unable to construct coherent conceptual knowledge.

While various technologies are often very helpful tools for teaching and learning, scientific visualizations, educational tutorials, and simulations software can be very extensive and complex for students' novices. However, Edelson and his colleagues (1995) suggested scaffolding students by creating a "front-end" of a simulation that allows students to learn basic steps and become comfortable with the upcoming task. These specifically designed supportive tools become cognitive tools or scaffolds (van Jooligen et al. 2007) within a simulation.

The scaffolding in computer-based interactive environments can take different forms.

For instance, Edelson, Pea, and Gomez researched three front-ends created for the *Collaborative Visualization* (*CoVis*) project: *Weather Visualizer*, *Climate Visualizer*, and *Greenhouse Visualizer*. The simulations employed a number of features to support learners' discussions, predictions, and their reflections on tasks. These features, as an example of integrated inquiry and collaboration tools, included adjustable manipulatives for creating observable changes, a notebook for making comments and keeping a journal of the observations, an opportunity to generate questions for supporting peer discussions, and support for creating plans and artifacts to demonstrate findings. Based on the created environments, researchers theorized that authentic settings engage students in scientific practice, collaboration tools enable this engagement, and social interaction enhances learning through the process of peer communication (Edelson, et al. 1995).

Just as an excessive use of interactive features and images, complexity of a computer-based environment can also disrupt cognitive learning processes. To minimize the complexity, such environment can be divided into smaller parts and use scaffolding tools such as templates to help learners move in small steps from an easier to a more difficult part of the learning environment.

IV. Methods for Supporting Learning in Computer-based Interactive Environments

Hmelo-Silver et al. (2006) pointed out that in order to make sense out of complex tasks, students must articulate their thinking, investigation process, and reflect on the learned material. Previous research findings of other investigators such as Barab and Duffy (2000), Mayer, Mautone, and Prothero (2004), van Joolingen and de Jong (1998), to name but a few, support this statement and show that there are various ways in which learners can be encouraged to reason and reflect on what they learn. One way to encourage students' thinking processes is cognitive tools embedded in computer-based technologies (Jonassen, 2000).

A study conducted by Lin and Lehman (1999) using a computer-based biology simulation in a college-level biology course showed that question-posing prompts are an effective tool for encouraging students' reflection. The 88 college students were divided in four groups and given a pre-test to establish their current knowledge of the subject.

One group of students worked with a simulation which included justification-prompts, second group worked with an emotion-focused version, third group used a simulation which required students' reflection on their experience working on the task, and third group had no prompts. As a result of the study, the researchers found that justification-prompts, in which students were prompted to give reasons for their actions, led to superior knowledge transfer as compared to two other types of prompts that encouraged students to explain rules or procedures and reflect on their emotions during their interactions with the simulation.

Another study using 38 college students examined how a computer-based environment can support reflective activities by means of visual feedback and verbal prompts (Saito & Miwa, 2005). The students were separated in two groups where one group worked with a simulation that required reflection from students on their actions while the other group did not have any prompts. The students' learning achievements were evaluated based on their scores on pre- and post- tests. The study showed that students who used the simulated environment had a better understanding of their actions when compared to a control group.

Cindy Hmelo (2001) investigated the effectiveness of encouraged reflection in a complex medical simulation using medical students. Simulated processes, from designing a trial to observing results, allowed students to gain practical experience in a significantly shortened time. The simulation helped students better understand interdependencies between multiple contributing variables in diagnosis and treatment processes (Hmelo, et al.). The researcher concluded that the experience through the simulation of inquiry process and encouraged reflection helped medical students construct better conceptual understanding of the subject.

Another study conducted by Wendy Adams et al. (2008) at Colorado University with 56 students was based on the *Physics Education Technology* (*PhET*) project. The focus of the study was to identify effective and ineffective characteristics of an educational simulation. The participants were college level students. Their interactions with the simulation were observed by the researchers. Additionally, students were asked to comment verbally on what they saw and expected from the simulation based on the use of manipulatives. After the testing follow-up questions on students' experience were asked. Based on their observations and interviews, the researchers found that students in physics classes learn best from simulations when the interaction is directed by students own questioning.

Learners' predictions of the result of their use of a simulation's manipulatives for adjusting a set of variables are also an important component in learning processes. Adams and her collaborators (2008) demonstrated that predictions based on previous knowledge, preconceptions and misconceptions, increased students' curiosity, engaged them in interaction and encouraged their deeper thinking. They conducted a qualitative study including 40 students who were asked a set of questions based on their interactions with a *PhET* simulated environment. The researchers also noted that predictions based on learners' misconceptions engaged students the most.

Wouter van Joolingen and Ton de Jong (1998) investigated how a cognitive tool can support students' building of hypothesis generation skills using *Hypothesis Scratchpad* simulation in a general science class. The simulation was designed as an interactive template. It was created to support cognitive processes by constraining students' expressions, content of hypotheses' space, and by providing manipulatives for a set of programmed variables. However, the findings of the study were not conclusive. While 12 study participants showed considerable improvement in creating more explicit and testable hypotheses, the post-test showed no significant change in learners' newly acquired skills. The researchers suggested that the students may not have spent enough time with the simulation to build a needed level of skills.

Furthermore, Manlove and Lazonder (2004) showed in their study that students in collaborative inquiry learning environments do not always stay on task. In their study the researchers used three groups of 13 participants each. The students' interactions with simulated models of the greenhouse effect were observed by the investigators. The results of the investigation's qualitative analysis showed that the learners required additional support to maintain engagement in the task and meaningful peer interaction. Two years later, Manlove, Lazonder, and de Jong (2006) conducted another study where the researchers implemented a scaffolding tool, the *Process Coordinator,* for regulating the students' activities. The study was designed to investigate the influence of guidance embedded in the simulation on students' learning outcomes. The scaffolding tool included a hierarchy of goals, hints, explanations, and a template for the task's final report. The participants of the study were divided into small groups. The students' learning outcomes were evaluated based on a created water tank model that demonstrated their understanding of the system dynamics. The results revealed that students in a treatment group not only outperformed students who use a similar simulation without a guiding scaffold, but they also spent less time discussing their understanding of the task and more on cognitive discussion of the learning task. However, there was no correlation found between high levels of the cognitive discourse and the quality of the model in the treatment group. This correlation was found in the control group only. After evaluating qualitative data, the researchers concluded that students, who collaborated in small groups of three peers, while participating in discussions did not share responsibilities equally and relied on the performance of only one member of a team to complete a model.

Lloyd Rieber (2004) explored across several studies (Rieber, 1996; Rieber, Chu, Tzeng, & Tribble 1996; Rieber & Noah 1997; Rieber, Noah, & Nolan, 1998) how adults and children learn scientific concepts from interactions with computer-based environment. He found during his investigations that interactions with an educational simulated environment and short instructional prompts, complemented by visual feedback, promoted reflective thinking and led learners to a deeper level of understanding (Rieber, 2004). The study also confirmed that even though the open-ended experiential simulations have many benefits, their interactivity does not *per se* contribute to students' understanding of scientific principles.

Gazit, Yair, and Chen (2005) investigated conceptual development of basic astronomy concepts by high-school students through their interaction with the dynamic 3D learning environment *Virtual Solar System* (*VSS*). 10 tenth-grade students participated in the study. The participants used the *VSS* for self-guided study of the Earth-Moon-Sun system. The students had to complete two tasks guided by observe-explain questions without any additional mentoring and one free exploration task. Students' learning achievements were evaluated based on open-ended conceptual questions administered as pre- and post- tests. The researchers did a comparative analysis of level of interaction (use of tools) with the number of correctly answered questions. The researchers noted an unequal frequency distribution in the use of tools, and overall low level of interaction, and a surprisingly high level of created misconceptions. The researchers concluded that "high interactive performance by students might not be sufficient for the development of scientific conceptual understanding" (Gazit, et al. 2005, p.468).

The results of the *VSS* investigation are directly related to, and support the rationale of, the proposed study which utilizes the guided inquiry-based learning environment aiming to constrain the users' actions. The rationale for using constraints, in the form of scaffolds for inquiring through students' predictions and reflections is to guide learners in a structured way through their own exploration while helping them form their own understanding of the parallax phenomenon.

Hence, the previous research works show that not every computer-based environment that allows users to interact with a simulated phenomenon possesses adequate educational value for supporting students' cognitive learning processes. Excessive unguided interactivity, an unbalanced amount of visual representation, and exceedingly complex environments disrupt learning processes making students lose their original goal and stop interacting in a structured way.

These potential and real problems can be avoided in employing a less demanding and more straight-forward educational tool – a computer-based tutor. Technology-supported tutoring systems in education many forms: from straight-forward drill and practice to adaptive and intelligent technology tutoring systems. Their role in learning communities were investigated from different angles such as usability (Edelson,Gordin,

& Pea, 1999), level of learners' control (Williams, 1993), and amongst others, user collaboration (King, 1998).

IV.I. Lerner-control and Program-control

The notion of using technology-based tutorials to facilitate learning processes can be traced at least back to the 1960s when researchers discussed positive and negative aspects of letting learners choose their own learning paths, by assuming control over the program. Michal Williams (1993) in providing historical background on learner-controlled computer-based environments, discussed differences between learning outcomes based on: users' prior knowledge and students' learning abilities (Mager & Clark, 1963), efficacy of the tutorials in higher-order learning of problem-solving and knowledge transfer (Campbell, 1964), and motivation (Campbell & Chapman, 1967).

In later years, efficacy of learner-controlled and program-controlled environments for learning achievements became an interesting subject for scientific discussions, not only have the discussions go on for decades, but they are still unsettled today.

A number of researchers argue superiority of learner-controlled computer-supported environments as they allow learners to be in charge of their learning, making educated choices (Ellermann & Free, 1990) based on their prior knowledge (Kopcha & Sullivan, 2007), interest (Lawless & Brown, 1997), motivation (Reigeluth & Stein, 1983), and desire to work at their own pace (Penland, 1979; Milheim, 1990).

For example, Kopcha and Sullivan (2007) investigated the effects of learners' prior knowledge and their preference to work in either a learner- or a program- controlled environment on their learning achievements in a math class. The researchers found that students with high prior knowledge who wanted to and were assigned to work in the program-controlled environment achieved better scores on the post-test. Students with low prior knowledge achieved higher scores on the post-test when their preferences of control were not matched. The investigators concluded that “matching learner preference to the type of program they receive is an effective strategy for high-prior-knowledge students but not for those with low prior knowledge” (p. 265).

Penland (1973), Hannafin (1984), and Hartley (1985) argued that learners would construct greater knowledge of the presented material if they have a choice of selecting their own path and are in a situation in which they have to constantly make decisions. A theory about the amount of mental effort invested in learning and the resulting outcome was proposed and investigated by Gabriel Solomon (1983, 1984) who suggested, “effort-demanding activities produce better recall, comprehension, and inference-making” (Solomon, 1983, p.44). He defines his *AIME* theory as the Amount of Invested Mental Effort, as “the number of non-automatic elaborations applied to a unit of material" (Solomon, et al). Mayer and Moreno (2003) in their cognitive load research, found that students show better understanding of multimedia explanation if it is broken up in segments in a learner-controlled environment. Reigeluth and Stein (1983) argued that learning is more effective in a learner-controlled environment if it is controlled by a motivated learner.

Other researchers found that when students received full control over their computer-supported learning they did not make a good use of it (Carrier, 1984; Johansen & Tennyson, 1983; Pollock & Sullivan, 1990; Reinking & Schreiner, 1985; Snow, 1980).

James Young (1996) examined how self-regulated learning strategies effected learners' performance in a learner-controlled and a program-controlled computer-based environment. The investigation was based on seventh-grade students who were divided into high and low levels of self-regulations. Results showed that students from the low level of self-regulated group performed poorly in the learner-controlled environment. In addition, it was found that that greater difference between performances of two groups was in a learner-controlled and less in a program-controlled environment. Thus, the researcher concluded that "program control seems to minimize the performance differences between low and high levels of self-regulated learning strategies" (p. 17).

Ross and Rakow (1981) compared two groups of students who learned math in a learner-controlled and a program controlled environment. 124 undergraduate students were divided into two groups and their results were compared twice: on an immediate post-test and a delayed post-test. The researchers found that students who worked in a program-controlled environment consistently outperformed students from a learner-control group on both post-tests.

Geraldine Gay (1986) conducted a study in which she compared learning achievements of high and low aptitude groups of learners who worked in learner-controlled and computer-controlled environments. 80 undergraduate students were randomly assigned into two treatment groups and their scores were compared on a post-test. Results revealed that high aptitude students received higher scores in both types of environments, while low aptitude students who worked in a computer-controlled environment outperformed low aptitude students who worked in a learner-controlled environment. Thus, the researcher suggested that "learners can be given more control if their prior understanding of a topic is high and should be provided with more structure if their prior understanding of a topic is low" (p. 225).

McNeil and Nelson (1991) in their ten-year long meta-investigation also concluded that students in computer-controlled environments tend to achieve higher results when they are compared to students working in learner-controlled environments.

As it was demonstrated above, while many researchers theorized that when learners exercised control over their learning experience they should achieve high learning results, these theories were not supported by empirical findings (Williams, 1993). Learning outcomes in both types of environments depend not only on the amount of control given to students but on their prior knowledge, learning aptitude and attitude, motivation, and self-efficacy, to name just few. It is very unlikely to find an optimal level of control for all students. It is likely, however, that even for one student level of control may vary from task to task (Lepper & Chabay, 1985).

IV.II. Scaffolding Tools

Scaffolding cognitive tools augment students' interactions with a range of computer-based interactive environments and help them continue working in a structured efficient manner. The scaffolding tools enhance learning processes and increase the educational value of the learning environments.

While computer-based interactive environments have the potential to be powerful learning tools that allow users to manipulate variables, the value of the environment partially depends on students' ability to draw on their prior knowledge and structure their actions so as to predict, infer, explain, imply, and justify their actions and outcomes. In other words, learners should be able to systematically apply logic (Russell & Munby, 1991) throughout the whole process of interaction. Such a process of reflection in inquiry learning stimulates reasoning in the students' learning process and deepens understanding of the learning subject.

As Colburn (2000) points out, if activities are too challenging, students cannot master the concepts; at the same time, if activities are too easy, students cannot achieve higher levels of thinking. Thus, the learning processes of complex concepts are often supported by scaffolding techniques and use tools to keep learners appropriately engaged in activities for achieving higher levels of performance (van Joolingen, Jong, & Dimitrakopoulou, 2007).

Inquiry-based learning assumes "extensive scaffolding and guidance to facilitate students' learning" (Hmelo-Silver, Duncan, & Chinn, 2006, p.99) unlike the unguided discovery approach. Students learn content and discipline-specific reasoning while being engaged in a collaborative investigation process.

Scaffolding is one of the methods of apprenticeship that is based on various forms of social interactions (Collins, Brown, & Newman, 1989; Dennen, 2002). Scaffolding facilitates learning where students may or may not succeed without an offered support (Hmelo & Guzdial, 1996). It helps students approach a complex task, makes the task less intimidating and keeps it within the students' reach (Hmelo-Silver, 2006). Scaffolding is meant to support learners' maximum development within their zone of proximal development (Vygotsky, 1978) and then slowly fade. One of the important features of the method is that it provides expert guidance to students on how and why the task should be done (Hmelo-Silver, et al.).

Over the years, researchers derived different types of scaffolding: (a) directive which is instructor structured, (b) supportive which is learner-negotiated (Lenski & Nierstheimer, 2002), (c) black-box which is a performance oriented approach that fades only when a learner becomes an expert, (d) glass-box which allows for performance and learning oriented approaches that utilize various kinds of support (Hmelo & Guzdial, et al.), (e) communicating process, (f) coaching, (g) hints, (h) visual animated representations, and (i) questioning prompts (Collins, et al.). This list represents just a few types of scaffolding emphasized by researchers in their studies.

In traditional face-to-face classes instructors can adapt to provide students with needed support, but in online environments scaffolding must be designed into the system and experience. Some types of scaffolding are used more than others in online learning environments and software-supported learning. Hmelo and Guzdial (1996) in their analysis of “software-realized scaffolding” (p.130), point out the successful use of prompts for self-explanations, intelligent agents, and collaborative environments. Therefore, students immersed in scientific learning via interactive environments, designed for supporting peer interactions, benefit from scaffolding features that support visual and social aid of students’ learning.

A scaffolding approach to conceptualized questioning was investigated by Hmelo and Day (1998), utilizing a computer-based simulation. The field-based study was conducted using 36 first-year medical students in a problem-based learning settings class. The participants were randomly assigned to one of two groups. The treatment group used a simulation for diagnosing the simulated patients’ illnesses while the control group used the traditional paper cases. The participants in the treatment group outperformed their colleagues in the control group on critical observations, 79% versus 66% respectively. The investigators found that the simulation encouraged students to "reflect on common themes" (p.163) while conceptualized questions helped the students to stay focused on the task.

Davis and Linn (2006) investigated influence of two different types of scaffolding prompts on students' knowledge integration which the authors describe as knowledge that "involves differentiating, integrating, and restructuring ideas" (Davis & Linn, 2006, p.820). The three studies were conducted among middle school students in science classes utilizing the *Knowledge Integration Environment* (*KIE*) software. The studies included comparison of self-monitoring and activity prompts and in-depth investigation of prompt responses and reflections. The researchers analyzed a total of 169 completed projects. In each study the students were divided in three groups (two treatment and one control groups). Researchers discovered that students who were scaffolded by self-monitoring prompts in which students were encouraged to plan and reflect on their actions achieved a high level of understanding on the topic. Students who were scaffolded by activity prompts designed to guide the inquiry process and asked students to identify and justify their actions provided less in-depth answers on the projects. Thus, the researchers concluded that not all types of the scaffolding strategies are equal in their merits and a mere involvement in a developed activity does not necessarily support students' building of deeper knowledge. In addition, the investigators suggested that self-monitoring prompts allow students to "make their own thinking more visible and explicit" (Davis & Linn, 2006, p.835).

Another learning environment *SMILE*, the *Supportive Multi-User Integrated Learning Environment*, provides learners with verbal prompts to help learners reflect on their experience (Kolodner, Owensby, & Guzdial, 2002). *SMILE* is a part of *Design Discussion Area* (*DDA*) system designed to help students to immerse into world of science, to help them to do experiments, and write reports and presentations in an acceptable scientific format including appropriate vocabulary and argumentation. *SMILE* includes hints, examples and templates to scaffold students reasoning and writing skills (Kolodner, et al., 2002). Owensby and Kolodner (2002) conducted a study with 47 eighth-grade students to examine how computer-based *Case Authoring Tool* (*CAS*), an additional component to *SMILE*, scaffolded students' ability to transfer newly constructed knowledge from one case to another and how the tool supported their collaborative work. During the experiment students were divided into comparison and experimental groups. The nine comparison groups included the 33 students who used a paper-based scaffolding template. The four experimental groups were comprised of the 14 students who used *CAS* for their work. Students' collaboration on assigned tasks and their in-class presentations were videotaped for qualitative inquiry analysis. The students' final performance was assessed based on their performance on a new case-based task. The results of the study suggested that the software supported groups developed "more sophisticated" (p.4) scientific reasoning skills.

As the studies demonstrated, there are a number of different types of scaffolding tools for supporting learners' cognitive processes: through collaboration, negotiation, reasoning, reflection, and visualization, to name a few. There are also many ways of implementing them within computer-based environments. However, not all scaffolds are suitable for all learning goals. Consequently, scaffolding tools should be specifically designed for computer-based environments' particular learning goals, as stated by *TSIL*. The scaffolds should also engage and challenge students' thinking, helping learners advance within their zone of proximal development. As was suggested by Lev Vygotsky (1978), social engagement and collaboration within learning activities are crucial components for successful achievement of maximum development.

IV.III. Social Component

The advocates of *TSIL* pay special attention to a social component of the theory and believe that by supporting conversations technologies can do even more and play a revolutionary role in learning (Duffy, Dueber, & Hawley, 1998; Edelson, Pea, & Gomez, 1995; van Joolingen, et al. 2007).

According to the constructivist view, meaningful learning is a cognitive process in which individuals make sense of the world in relation to the knowledge which they have already constructed. This sense-making process involves active negotiation and consensus building (Fraser, 1998) where meaning is a negotiated construct, built in a process of social interaction (Barab & Duffy, 2000). Thus, learning takes place in a social context through the learners' interaction, observed and/or imitated behavior (Bandura, 1977), and perception of these actions (Bandura, 1986). Learners' attention, memory of prior experience, and motivation contribute to the social interaction and social learning outcomes. As a result, social interaction is a key factor of learning. As Gunawardena (1995) points out, there is no learning without interaction.

It is widely believed that students learn better within an authentic context (Dewey, 1938; Vygotsky, 1997). Jean Lave (1988) also argues that learning is situated and our actions are embedded in the concrete situation in which they occur (Anderson, 1996). Through interaction learners become involved in a "community of practice" (Wenger, 1998). In the course of collaboration within the community, learners observe and model behavior, acquiring new skills and constructing knowledge. Edelson, Pea, & Gomez (1995) believe that a scientific community can be treated as a community of practice which is shaped by shared language, activities, and values.

However, while a group of students working together helping each other to prepare for classes and collaborating on course assignments during a semester can be considered a community of practice, brief in-class interactions with peers is a different aspect of social participation. Short-time interactions are important units of social collaboration in classroom settings. Change in understanding a phenomenon, restructuring knowledge and mental models, as a result of short-time span interactions, is a microdevelopment approach for studying a process of learning in real-time (Granott & Parziale, 2002). Short-time span interactions include learners’ collaboration, interaction and/or use of tools. These short-time interactions give a good idea to social scientists about how learning takes place in real time (Gazit, et al., 2005). Eric Mazur (1997) suggested that short interactions among peers in the course of discussions of their answers to questions posed in class, produce answers of a higher level. Task-oriented short-time interactions among peers engage students in sharing their knowledge and personal understanding of a subject thus encouraging learners’ deeper thinking processes.

Sharing thoughts, testing ideas, and discussing conclusions cannot be done individually. Collaborative work encourages students’ thinking process: it makes them doubt their thoughts, rethink their decisions, double-check the evidence, redraw conclusions, recognize gaps, and create connections between disconnected pieces of knowledge (Edelson, et al., 1995). The positive aspect of communication between peers during a collaborative work is a well-researched topic in social sciences. Eric Mazur (1997), for example, suggested use of peer instructions in Physics classes for helping students building conceptual knowledge. His five-year research shows that, on average, students gained 29% on the tests after participating in peer instruction activities.

Based on peers’ learning tasks, assigned roles, and their collaboration and interactions within a social context along with the nature of the study (Palincsar, Brown, & Martin, 1987) these interactions can be referred to as peer interaction, peer instruction, peer-to-peer, and think-pair-share, among others. While each of these terms has more or less subtle differences from others, peer interaction is a general term that encompasses these types of social interactions during students’ collaboration. Peer interaction assumes discourse participation of two or more of more or less equal parties (Guiller, Durndell, & Ross, 2008; Gupta, 2008; LaPointe, Gunawardena, 2004; Schwarz & Linchevski, 2007). Peer instruction (Mazur, 1997) assumes that peers not only share their ideas but instruct each other on a subject or a given task. Think-pair-share (Hernández-Leo, Villasclaras-Fernández, Asensio-Pérez, & Dimitriadis, 2006) often refers to a problem-solving discussion between two peers of the same level of knowledge on the discussed topic. Peer-to-peer (Bostrom, Gupta, & Hill, 2008) assumes negotiation of meaning between two or more peers. Since all these titles refer to the same idea of peers collaborating on a given task, discussing it, making decisions, and being involved in distributed knowledge activity, this proposal refers to this type of collaboration as peer interaction.

Duffy et al. (1998) point out that working collaboratively during an inquiry process is not the same as working individually; hence, environments created by learning technologies must support the social aspect of learning. The researchers also distinguish between two types of peer collaboration: conversation and issue-based discussion. In the first case, it is a “me”- centered environment where peers share and assess each other’s knowledge and opinions on a problem. This process is more opinionated than inquiry-based. Later students move to a discussion where they focus the collaborative effort on the task: developing plans, building hypotheses, and working on finding a solution (Duffy, et al., 1998). Thus, the researchers believe that technologies must support students’ inquiry through not only interactive tools but by supporting the students’ collaborative processes. The same idea is reflected by *TSIL*. It differentiates between particular tools designed for engaging students in social and computer-based interactions. Tools designed for interactions aim to enhance learners’ performance, while tools designed to aid students’ advancement with their inquiry to a higher level of knowledge construction when support is no longer offered, are often social interaction supporting tools (van Joolingen, et al., 2007).

Consequently, educational computer-based interactive environments built around the concept of constructivist learning or learning by doing (Gibbs, 1988; Shank, Berman, & Macperson, 1999) provide learners with authentic settings of simulated realistic surroundings where students are enabled to create groups and work on assigned tasks together. In dynamic computer-mediated learning environments users interact with each other using a number of specifically designed tools. Through the social interaction students discuss their tasks and assignments, plan next moves, share opinions, explain their findings, and argue their positions.

These short-term learning communities create unique learning cells where students become either equal participants or masters and apprentices. Students learn from each other while collaborating on assigned tasks, building knowledge together through the process of negotiation of meaning, and reflecting on their own actions (Kolodner, Owensby, & Guzdial, 2002). In this case, learners are involved in processes of reflecting and reasoning while collaborating with peers; they learn from each other in a form of peer tutoring (King, 1998), a form of peer interaction. The success of such interaction between peers was recorded by Palincsar, et al. (1987). The researchers conducted a study in which they compared achievements of 24 high school students divided into tutors and theirs tutees. The tutors were instructed by teachers on a selected subject; when mastery was achieved the tutors went on working with their peers helping them to gain the same understanding of the subject. The students' understanding of the subject was evaluated by the class instructor at the end of the semester. The results of the study showed that tutees' gain was comparable to the gain of their peer tutors. Alison King (1998) also studied building of new knowledge as an outcome of collaborative effort of tutoring pairs comprised of high school students of similar age and abilities. The investigation was based on the author's Guided Peer Questioning approach in which peers explained to each other concepts following a set of the designed questions. The researcher observed and recorded the students' interactions. Students' achievements were measured on the number of correct answers given while they were working with the simulation. The results of her study confirmed the previous findings and suggested that negotiation of meaning during peer interaction facilitates construction of new knowledge.

Another study that emphasizes the importance of peer interaction was conducted by Webb (1982) among 77 high school students. The researcher examined the correlation between students’ interactions and their achievements. Among other conclusions, the researcher found that students who gave explanations to other students during peer interaction, scored higher on the achievement test as compared to students who seldom gave explanations. Webb attributes this finding the mechanism of cognitive rehearsal, where students verbalize material, and to cognitive reconstruction, where students reorganize material for clear presentation, while explaining it to others.

A computer-supported 3D virtual learning environment is used by the University of Colorado *Business Computing Skills* (*BCOR*) course. The course uses the environment to assist students in building social and business skills by means of virtual interactions as well as encouraging students’ collaborative problem solving. The designers of the course indicate that visual environment resources and real-life communications support students’ collaboration and help to achieve course goals (Dickey, 2005).

Nevertheless, not all research investigations corroborate the previous studies’ results. Ge and Land (2003) conducted a study in which the researchers examined the use of question prompts and peer interactions to scaffold 117 students on ill-structured problem-solving tasks. The researchers found that students who were scaffolded by the question prompts performed significantly better on the tasks. However, the researchers’ prediction that the students working with their peers would outperform students working individually was not confirmed by the study’s results. Furthermore, the researchers’ other hypothesis that students who worked with peers and were scaffolded by the question prompts would demonstrate even higher levels of solving problems was not supported by statistical analysis of the study. These unexpected results contradict previous positive correlations between peer interactions and outcomes found by Palincsar, et al., King, and Webb. At the same time, qualitative findings of the Ge and Land study revealed a certain pattern among students' peer interactions: students spent more time discussing problems in the beginning of the collaborative work and were barely involved in discourse by the end of the task. The researchers suggested that this behavior of students not being completely engaged in the task along with the small sample size might explain the unexpected results. Ge and Land also concluded that while question prompts could be helpful to guide and keep peers and individuals focused on the task, they might not provide enough support for students to generate their own questions and to clarify each others' understanding (Ge & Land, 2003). The investigators suggested that to "maximize interactions thought interpretation, elaboration, explanation, negotiation, and argumentation" (p.35) other types of prompts are needed.

Another investigation revealed similar findings. Pilkington and Parker-Jones (1996) conducted a study in which medical students worked with simulations of calcium homeostasis while developing conceptual understandings of a very complex system – the human organism. Forty-two participants of the study had to predict how calcium would be controlled by the organism based on several variables, simulated data, and a feedback loop. Students were divided into two groups of 22 and 20 participants. One group consisted of ten pairs working together on the simulation and the other group consisted of 22 students working individually. All students took a pre-test before they started working with the simulation. After the participant finished their work, they were given a post-test. Students in both groups were constantly prompted to reflect on their actions by a set of questions. The results of the study showed that students who spent more time interacting with the simulation gained a deeper understanding compared to their classmates who discussed questions in a peer interaction form.

The results of the two later studies which support the Duffy et al. (1998) research suggest that two types of collaboration: conversation and issue-based discussion among peers have a significant influence on the inquiry process and learning outcomes. Furthermore, the researchers’ conclusions also support the *TSIL* concept of the need for creating and employing particular cognitive tools for particular activities and context.

As has been demonstrated in most cases, peer interactions engage students in deeper thinking and constructing better understanding of a subject. Learners gain the most through interacting with peers when the interactions are encouraged and supported by appropriate tools.

However, collaboration among peers does not always works the way it was intended by instructors and researchers. The social component of education research is always a current topic for discussions.

IV.IV. Peer Interaction in Learner- and Program- Controlled Environments

A number of researchers investigated effect of cooperative learning of peers working in small groups and in pairs and their interactions with computer-supported learner- and program- controlled environments. While some researchers advocate benefits of use of mixed heterogeneous grouping in which low-ability or low aptitude students can benefit from working with high-ability or high aptitude students (Heift & Nicholson, 2002; Mitrovic, 2003; Webb, 1982), other researchers recognize this type of mismatch as a detriment factor for both types of learners. The scientists argue that in heterogeneous groups able students would learn less as a result of distraction while less able students would not improve their knowledge but possibly achieve higher scores at the expense of more able students (Hooper, 1992). To investigate these inconclusive findings, Simon Hooper (1992) set up a study in which he compared learning in small groups with individual learning in homogeneous and heterogeneous groups in learner-controlled and program-controlled environments. 115 fifth and sixth grade students were identified as high- or average- level ability learners and were randomly assigned to work with peers or independently. Peer groups were divided into homogeneous and heterogeneous subgroups. The researcher found that high-ability students who worked in a homogeneous subgroup achieved the highest results while average-ability students who were grouped in a homogeneous subgroup achieved the least. In addition, Hooper pointed out a significant and positive correlation between students' helping behavior and their achievements which allow investigator to suggest a connection with amount of invested mental effort (Solomon, 1984).

Singhanayok and Hooper (1998) investigated the effect of peer interaction in small groups of high and low achieving students in learner- or program-controlled computer-based environments. 92 sixth-grade students were divided into high- and low-achieving based on their scores on a prior administered test. Then students were randomly assigned to work independently and in small peer groups. Results revealed that students who worked with peers performed better regardless whether they were identified as high or low achievers and a type of environment they worked with: learner- or program- controlled. These students also had better attitudes toward grouping. Also, students who worked with peers showed improvements in scores both on immediate tests, and delayed post-tests. High-achieving students performed better in the learner-controlled environment while low-achieving students succeeded in the program-controlled environment. The investigators concluded that cooperative (peer) learning: "provides beneficial effects, and implies a need for software designers to adapt computer-based instruction for cooperative learning to the different learning styles of high-and low-achieving students" (p. 17).

Hooper, Temiyakarn and Williams (1991) conducted a study examining a role of cooperative learning in small groups of students on their learning achievements in a learner-controlled and in a program-controlled environment. 175 fourth graders were identified as high and average ability pupil and were randomly assigned into treatment groups to work with peers or independently in two types of computer-supported environments. There were several findings of this study: (1) students who worked with peers showed higher learning achievements and better attitudes toward their learning experience; (2) students in the program-controlled environment completed more exercises than the students from the learner-controlled environment; and (3) there was no difference in effect of type of environment on students' leaning achievements and attitudes.

Crooks, Klein, Savenye, and Leader (1998) investigated effects of cooperative learning with peers and independent individual learning in a learner-controlled computer-based environment. The investigators examined undergraduate students' achievements, use of environment, attitudes and social interaction. The students were divided into groups to work independently or with peers. The environments also allowed users to choose different levels of control. The results of the investigation showed that students who worked with peers spent more time on practice and other interactive features of the environment as independently working students. However, these interactions did not reflect on students achievements. The investigators concluded that "the achievement benefits of cooperative learning found in previous research may not apply to situations in which mature students are provided with an instructional environment with many learner-controlled options" (Crook et al., p. 223).

As it was demonstrated above, there is no one well-fitting solution ready to be applied to different types of tutorials. Thus, this study investigated peer interaction as it was supported by *SPIRUT*'s built-in scaffolding tools which encouraged students' discussions and cooperative work.

V. Developing Abstract Thinking and Transfer of Knowledge

Transfer of knowledge is the ability of a learner to apply constructed knowledge to a situation different from the one in which this knowledge was constructed. It is an ability "to extend the knowledge and skills one has developed beyond the limited contexts in which they were acquired" (Pellegrino, Chudowsky & Glaser, 2001). It should be noted that transfer of knowledge is closely related and often requires abstract thinking.

Abstract thinking can be defined as a style of thinking in which understood concepts and ideas are applied in problem solving, new ideas and surroundings, and/or translated in other cognitive domains such as mathematical applications (Clements & Sarama, 2004). Thus, the ability to generalize from a concrete experience is an important part of abstract thinking.

However, if learners cannot connect one or more aspects of a phenomenon or several concepts into another more complex one, they cannot build necessary abstract thinking skills and transfer their knowledge to a new situation.

There has been substantial research done in the field of computer-based interactive environments for teaching various disciplines that suggests that these environments are especially effective for teaching difficult and abstract subjects where students often have to understand the causal relationship among multiple variables and build a conceptual understanding of a subject. Results have been inconsistent. While some social scientists report a successful use of computer-based tools, others report existing problems.

Stieff and Wilensky (2003) used a simulation of chemical equilibrium in an attempt to enhance students' reasoning and conceptual understanding of interconnectivity in the system. The researchers reported that after interacting with the simulation "each student came to depend less on algorithms and rote facts and more on conceptual approaches to problem solving and answer justification" (p. 299).

Warnakulasooriya and Pritchard (2005) investigated efficacy of a web-based tutor, MasteringPhysics, on students' problem-solving skills taking into account time spent on tasks and accuracy of the provided answers. The study was conducted at

Massachusetts Institute of Technology (MIT) among 400 students enrolled in the *Introductory Newtonian Mechanics* course. The investigators found that students who used the computer-tutoring help in a form of hints, descriptive text, and feedback significantly outperformed a group of students who solved the problems without any additional help. In addition, the researchers discovered that after using hints and feedback, over 93% of the students ultimately found the correct solutions to the presented problems.

On the other hand, Frederiksen and White (2002) conducted a study where they demonstrated how principles of electricity can be successfully understood by learners through interaction with simulated models in science classes. The researchers pointed out that while a number of simulations (in this case) were used effectively in science classes to demonstrate scientific phenomena, teachers rarely brought students attention to "conceptual linkage" (p. 69) among the used simulations.

Furthermore, students often have difficulties translating visual images to other applications such as applying an abstract concept to a new situation. In research of the impact of a scientific educational simulation on students' understanding of Kepler's Third Law (Ruzhitskaya & Speck, 2008) using a *CLEA Project* simulation, we discovered that the simulation did not improve students' mathematical abilities. The study was conducted with 26 non-science major college students whose learning achievements were measured using pre- and post- tests created by the software developers. Even though students had a significant gain on conceptual questions they had no gain on a question requiring a mathematical application. The results of the study suggested that regardless of the students' high level of engagement working with the simulation, the disconnection between visual representations and the ability to perform a simple mathematical task were not bridged by the simulation's tools.

Consequently, it can be inferred that the use of visual representations, direct instructions, and constrictions provided by a computer-based interactive environment are not enough for developing the mathematical application of abstract thinking. In order to help learners translate visual representations into mathematical solutions, new scaffolds and ways of implementations need to be developed and investigated.

VI. Conclusion

Following *TSIL*, technologies can be used as cognitive tools in support of science education when the following components are included: the identification of a motivational context, the selection and sequencing of activities, the design of investigation tools, and the creation of process supports. The theory suggests that the motivating context should be identified early in the development process and should focus on meaningful, controversial, and open scientific issues. The selection of activities and their sequence should support students' interest in doing the investigation using appropriate levels of interactivity and available tools as well as meeting increasing levels of difficulties bridged by an appropriate scaffolding design. The design of investigative tools should be based on the careful selection of manipulatives for specific task and learning outcomes goals.

The current study expanded the findings of the previous studies and investigated the effect of scaffolding - in a form of questioning, feedback and a program-control option of *SPIRUT* - and peer interaction on learning a scientific concept, improving mathematical skills and transfer of the constructed knowledge.

This investigation examined scaffolds as cognitive tools for engaging students in processes of prediction, manipulation of variables, and reflection in the inquiry-based tutorial. The study aimed to find out how peers interaction and guided technology-supported inquiry learning influence students’ learning achievement of scientific concepts.

CHAPTER 3

METHODOLOGY

I. Introduction

The proposed study utilizes a mixed method of inferential statistics and qualitative methods to examine the impact of inquiry based methods and peer interaction on the students' construction of conceptual knowledge.

The study investigates how the use of inquiry-based tutoring tools, such as computer-based and paper-based tutorials, program-controlled and learner-controlled environments, and working with peers or individually contributes to students' learning the concept of stellar parallax. The investigation compares students' learning outcomes as the result of the peer and individual interactions with two types of tutorials as well as students' interaction in two types of computer-based environments (learner-controlled and program-controlled) to investigate whether students construct greater knowledge of the scientific construct while engaged in activities using one or the other tutorials.

In addition, the study examines how these methods improve students' mathematical skills and aids students' transfer of knowledge - transformation from visual representations of the scientific construct depicted in the tutorials and guided mathematical exercises into their ability to apply this knowledge through engaging of abstract thinking in unfamiliar situations such as finding distances anywhere in the galaxy by using the stellar parallax method.

II. Educational Tools

To investigate efficacy of a paper-based and computer-based tutoring tools, *Lecture-tutorials for Introductory Astronomy*, (*LTs*) (Prather, Slater, Adams, & Brissenden, 2008), a workbook of astronomy tutorials, and *Stellar Parallax Interactive Restricted and Unrestricted Tutorial* (*SPIRUT*), developed by the author, were used in the study.

II.I. Lecture-Tutorials (*LTs*)

Lecture-tutorials (*LTs*) is a well-recognized workbook developed for introductory astronomy courses for non-science major students by Edward Prather, Timothy Slater, Jeffrey Adams and Gina Brissenden (2008), (see Appendix B). The tutorials are extensively used in universities and colleges across the country (Prather, et al., 2009).

LTs is a set of paper-based collaborative learning activities (Prather, Rudolph & Brissenden, 2009) on a range of astronomy topics. This inquiry-based tutoring tool is based on Socratic questioning in which students engage in a science process of recognizing facts, setting up predictions, and drawing conclusions. The activities, in *LTs*, are set up in a form of questions and hypothetical dialogs between two or three students in increasing level of difficulty with minimum explanations so the students can construct their own knowledge of scientific constructs based on given evidence.

LTs is specifically designed to be used as a learning tool during peer interaction and “to be completed by pairs of students in 10-20 minutes working together in lecture-hall settings after having heard a short lecture on the relevant topic” (Prather, Rudolph, &

Brissenden, 2009). The role of an instructor in these activities is as a mere facilitator of the discussions. It is a premise of the workbook that students construct greater knowledge by means of their own collaborative work with peers.

II.II. Science Inquiry, Computer-Based Interactive Tutorial - *SPIRUT*

To investigate learning processes facilitated by a computer-based integrative interactive environment a computer-based learning tool, *Stellar Parallax Interactive Restricted and Unrestricted Tutorial* (*SPIRUT*), was developed. *SPIRUT* is an introductory component of a computer-based laboratory module of the *Stellar Properties Laboratory* project (*SPLab*). *SPLab* was developed to help students practice and deepen their understanding of several aspects of the stars such as distances, apparent and absolute visual magnitudes, luminosities, temperatures, and radii. *SPIRUT* was designed to explicate the concept of the stellar parallax and its application for finding distances to stars. The goal of the tutorial is to allow students to construct their own understanding of the concept through the use of tutorial’s tools for conducting observations, collecting data, and converting units of measurements.

SPIRUT was designed to enhance students’ knowledge construction processes of an astronomy concept of the stellar parallax. *SPIRUT* was also aimed at enhancing students’ learning experience through interactive exercises and visualizations. *SPIRUT* uses a set of tools that according to the technology-supported inquiry learning theory should facilitate students’ learning process without overwhelming them and creating unnecessary destructions. Some education researchers (Jonassen, 2000; Berson, 1996), while accepting an educational role of computer-based tutorials as one of the forms of computer-based mind tools, also express their concern about use of computer-based tutorials based on the tutorials' two common restrictions: impossibility to pre-program a tutorial to be able to react to each student input, and precluding students from constructing their own meaning but rather "map a single interpretation of the world onto what is important reflect on and assess what they already know [resulting in] inert knowledge" (Jonassen, 2000, p.6).

SPIRUT was created keeping in mind these concerns with a goal of minimizing the mentioned constraints. The tutorial includes a set of scaffolding tools that lead students through a set of exercises from predictions, observations, reflection and data collection to calculating a distance to a star. Since students are involved in learning a new concept, there is a minimal risk that students will use their preconceived notions and since they are constantly involved in a process of predictions and reflections, the second concern is minimal risk as well.

Stellar Parallax Interactive Restricted and Unrestricted Tutorial, as is suggested by the title, includes two versions: restricted or program-controlled and unrestricted or learner-controlled. In the program-based version of the tutorial, students' interactions and their progress in the environment are controlled by the program. *SPIRUT* is an adaptive program that responses to students' actions and answers given to program-generated questions by selecting a next step and directing learners from scene to scene. In the learner-controlled version of *SPIRUT*, students are allowed to move freely from one part to another, from one scene to another skipping or repeating, answering questions, and engaging in interactive exercises at their will.

II.II.I. Structure of SPIRUT

The tutorial consists of four parts: I simple geometry, II units' conversion, III trigonometric parallax, and IV stellar parallax. Only the last part will be described in detail here, the other parts are supporting sections of the tutorial to which students are directed to refresh their knowledge, or brush up on basic geometry and unit mastery. Each part starts with a set of prediction questions directly addressed to students by their names. Based on their answers to the questions and subsequent exercises, students are, in the case of the program-controlled version, either taken to the next step of that part of the tutorial or they are taken back to an appropriate part of the tutorial to review. In the case of the learner-controlled version, students are not taken back, but rather a strong hint is displayed that the students might want to move to a more appropriate part of the tutorial. *SPIRUT* is constructed in such a way that when student comes back to the same exercises and gets it wrong again, that they will be directed (hinted in the learner-controlled version) in yet another part of the tutorial. If they get it wrong again, the correct answer is given, and the students can move on.

In the following more detailed description the program-controlled version of the tutorial is assumed. For the learner-controlled version, simply substitute every forced redirection for a hint to the students to redirect themselves.

Preamble, started on Tue Jan 11 14:43:24 2011

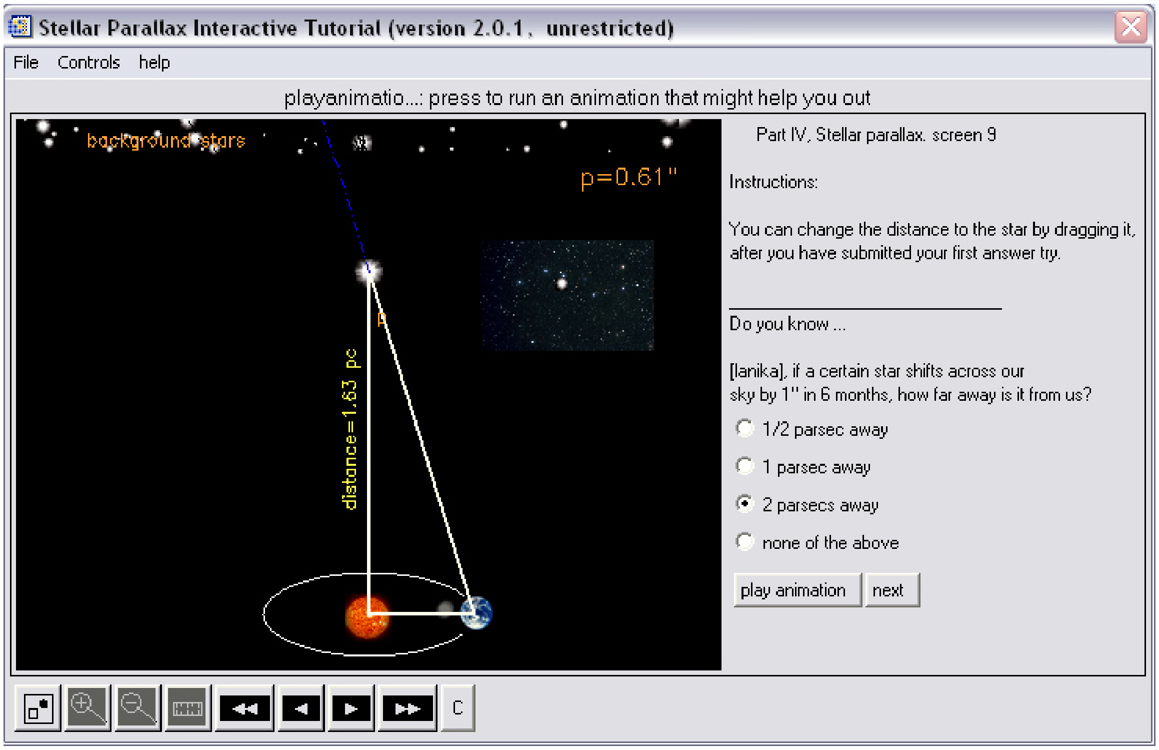
To get started please give your name(s) and press submit when done

users entered so far: 0

Figure 3.1: The welcome screen for SPIRUT, where students enter their names.

SPIRUT starts with the last, most advanced part of the tutorial, in which students are required to perform observations and calculate distance to Sirius, a star in our sky. The first scene (Fig. 3.1) of the part is an introduction, and the second scene offers a short-answer question asking about a baseline used to determine the Sirius' parallax which allows students to choose their starting point based on what they think they know about the concept. Students can answer the question or select "I would like to refresh my knowledge first" button in which case they are taken to the beginning of part III, Trigonometric Parallax, to refresh their memory on terminology involved in the concept and relationship between the parallax, baseline and distances to distant objects. If students overestimated their knowledge and provided a wrong answer, the tutorial takes them back to the same beginning of part III. Students who answered the question correctly advance to scene 3 where they are interact with the tutorial and change the distance to a star while observing the concomitant change of its parallax and parallax angle. Here students are asked to predict how long will it take for a star to shift from the rightmost to the leftmost position in the sky. After selecting a correct answer out of five choices, students advance to the next screen. Students who picked a wrong option are given feedback and asked to watch an animation showing the earth moving in its orbit and how its motion affects the apparent shift of the star in the sky. After that, students can resubmit their answer and then are taken to the next scene.

The next three screens contain a set of prediction questions asking students to envision a number of situations such as change of baseline length and observing a star's shift from Mars's orbit, or observing a star that is located not right above the plane of the solar system but at a different angle to our line of sight. Every time when students submit an incorrect answer they are provided with feedback in a form of a hint and asked to watch an animation specific to the scene, or a visualization of a described situation, that allows students to rethink their position and resubmit their answers. Students who provide correct answers are taken to the next step - explanation of the relationship between the earth's orbit, parallax angle and the parallax shift of a star. The explanation includes a mathematical representation of the relationship as expressed by formula $d=1/p$ and this explanation is supported with an animation. These two screens of part IV also includes an animation showing how arcseconds and degrees are linked together.



playanimatio...: press to run an animation that might help you out

Part IV, Stellar parallax. screen 9

Instructions:

You can change the distance to the star by dragging it, after you have submitted your first answer try.

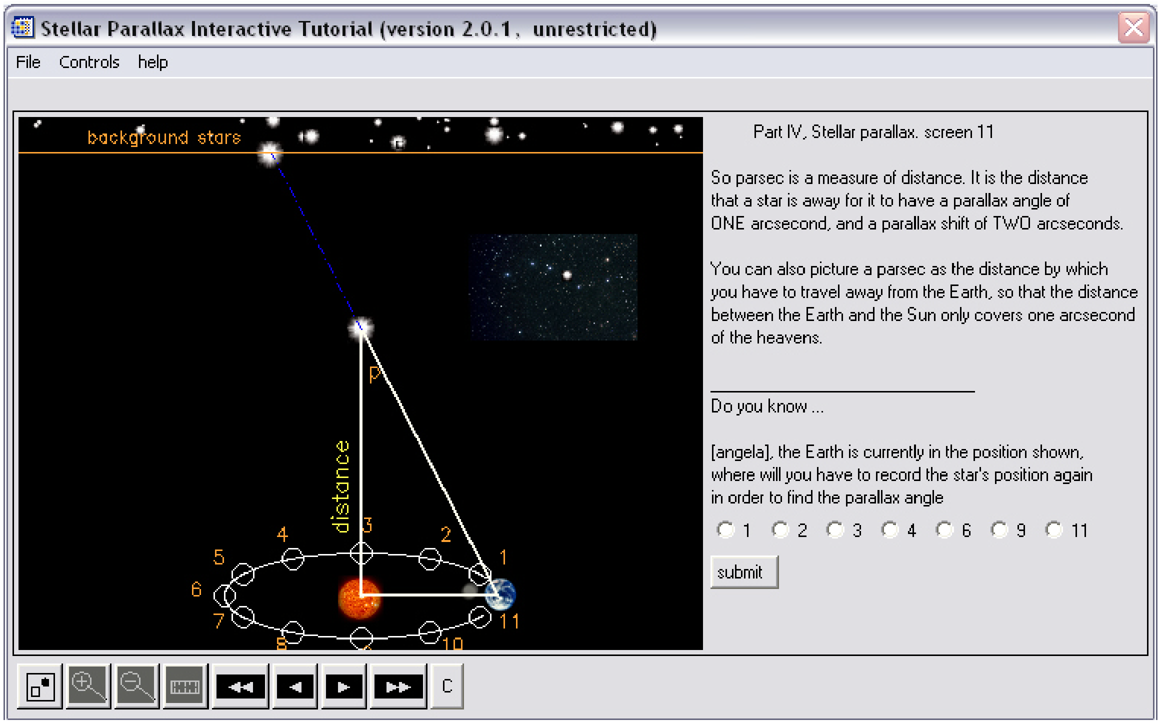
Do you know ...

[lanika], if a certain star shifts across our sky by 1" in 6 months, how far away is it from us?

○ 1/2 parsec away
○ 1 parsec away
● 2 parsecs away
○ none of the above

Figure 3.2: A screen snapshot of the part in the tutorial where students can adjust the distance to a star, and witness the accompanying changes in parallax angle.

After reading explanations and watching animations, students are asked a new set of a multiple-choice prediction questions. This time, the questions call not only for conceptual understanding but also for understanding of the relationship when numerical values are involved. For instance, 'If a certain star shifts across our sky by one arcsecond in 6 months, how far away is it from us?' Students, who understand the concept, should be able to do a simple calculation and pick the right answer of two parsecs away. Students, who selected a wrong answer, are asked to use a cursor to adjust the distance from Earth to the star (Fig. 3.2) and watch how the values of parallax angle change with the changing distance. Note that the question is asking about parallax, not parallax angle: the interactive exercises hint students to fathom the relationship between parallax and parallax angle while it does not give a direct answer to the students.



Part IV, Stellar parallax. screen 11

So parsec is a measure of distance. It is the distance that a star is away for it to have a parallax angle of ONE arcsecond, and a parallax shift of TWO arcseconds.

You can also picture a parsec as the distance by which you have to travel away from the Earth, so that the distance between the Earth and the Sun only covers one arcsecond of the heavens.

Do you know ...

[angela], the Earth is currently in the position shown, where will you have to record the star's position again in order to find the parallax angle

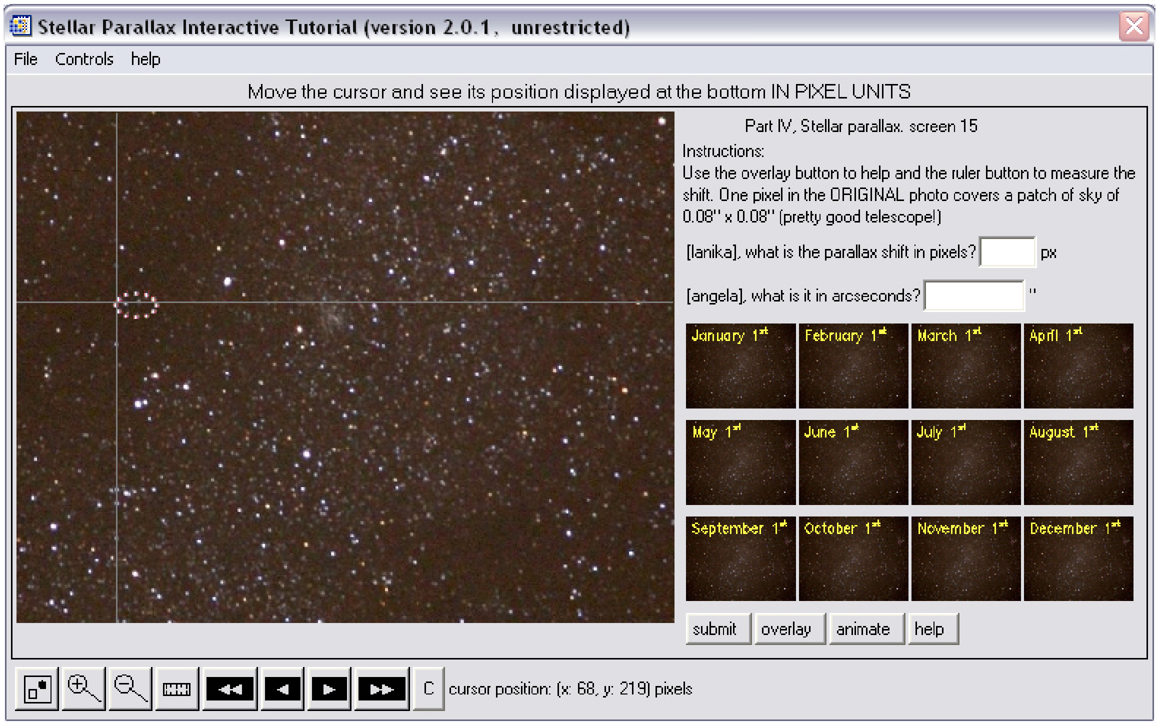
○ 1 ○ 2 ○ 3 ○ 4 ○ 6 ○ 9 ○ 11

Figure 3.3: In this part of the tutorial, students will have to correctly predict at which orbital positions the parallax shift is measured.

The next set of questions verifies that students understand that measurements of the parallax shift must be taken 6 months apart. Based on the position of Earth in its orbit when the first position of a star was recorded, students are asked to select the next position for the second recording of the star's position (Fig. 3.3, screen 11). If students selected the wrong answers, they would receive a hint and feedback in a form of an animation where they can observe how the position of a star changes in the sky as the Earth revolves around the Sun.

After these exercises students are asked to reflect on what they have learned. Students are asked to explain in two to three sentences why they have to take measurements of the stellar parallax six months apart.

After students answered the three sets of prediction questions, read explanations of the stellar parallax, participated in the interactive exercises, watched animations, and concluded their work with reflection on the relationship between the earth's orbit and the stellar parallax, it is time to apply this conceptual knowledge. Over next three screens students are involved in using real photographs of a patch of the sky with a superimposed simulated parallax of a star (Fig. 3.4). In this final section of part IV of the tutorial, students are actual astronomers who have to overlay 12 photographs in order to find a moving star on the background of more distant stars.



Stellar Parallax Interactive Tutorial (version 2.0.1, unrestricted)

Move the cursor and see its position displayed at the bottom IN PIXEL UNITS

Part IV, Stellar parallax. screen 15

Instructions:
Use the overlay button to help and the ruler button to measure the shift. One pixel in the ORIGINAL photo covers a patch of sky of 0.08'' x 0.08'' (pretty good telescope!)

[lanika], what is the parallax shift in pixels? ____ px

[angela], what is it in arcseconds? ____ ''

cursor position: (x: 68, y: 219) pixels

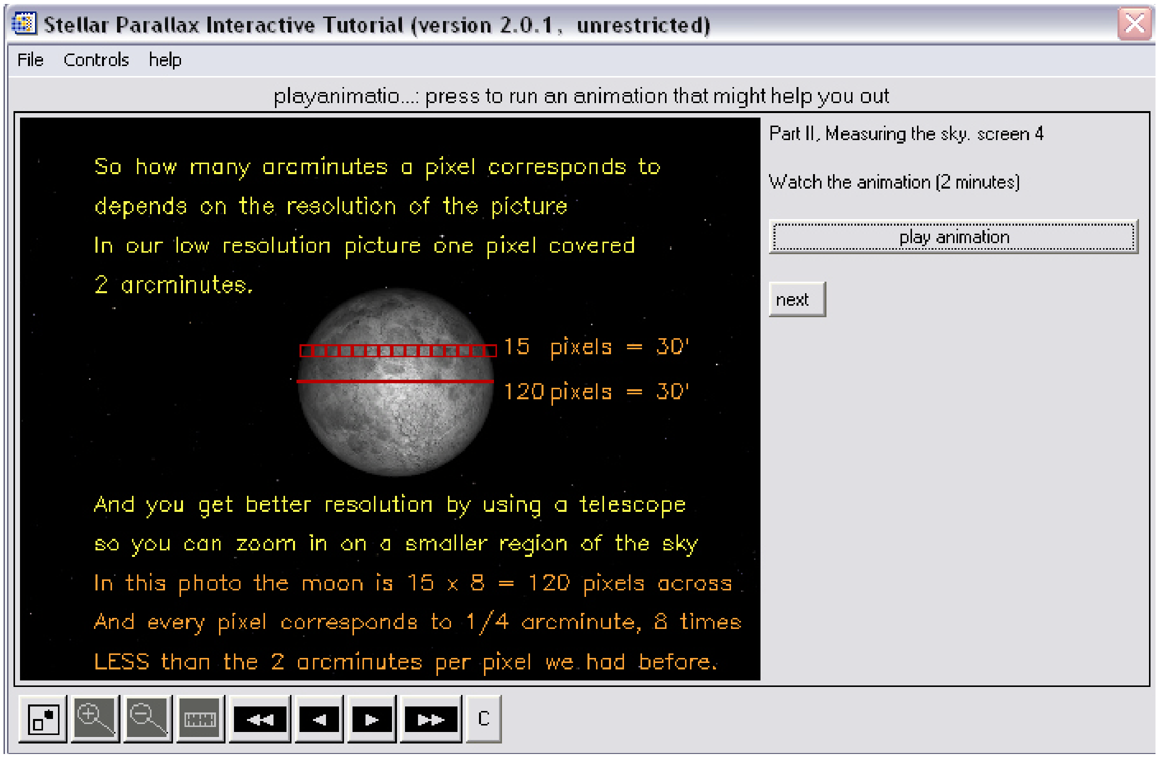
Figure 3.4: In this part of the tutorial students act like real astronomers, overlaying pictures, measuring the parallax shift and calculating the distance to the star.

Next they have to measure its shift on a computer screen in pixels, convert it in arcseconds, and use the parallax formula to calculate the distance to the star (Fig. 3.4).

This concluding segment of the tutorial ties all the previous exercises and animations together and takes students to yet another level of detail in order to construct knowledge on how to find and use the stellar parallax, and then demonstrate its use in practice, which involves an understanding of how to deal with varying resolutions of observing instruments and conversion of units. To help students to find the distance to a star, the section includes an animated feedback on how to convert units from pixels to arcseconds, the importance of a photograph's resolution and how to account for an applied digital zoom (Fig. 3.5).

After students successfully calculated the distance to the star, they are asked to explain in two to three well-structured sentences how astronomers find distances to the stars using relationships between parallax shift, parallax angle, the Earth's orbit and its distance to the Sun.

SPIRUT is programmed to address directly each student who is working in a peer group. This assures that each student is responsible for negotiating and providing the answers and thus the program stimulates collaboration between the peers. For instance, the questions on screen 15 (Fig. 3.4) are directed to both students. Other examples would include asking one student: "[Name1], the Earth is currently in the position shown. Where will you have to record the star's position again in order to find the parallax angle?" The second student is offered a different, yet conceptually similar question: "[Name2], if you want to find the distance to a star, which is right above the plane of the solar system, you should record its positions in the sky when the Earth is at positions…"



playanimatio...: press to run an animation that might help you out

Part II, Measuring the sky. screen 4

Watch the animation (2 minutes)

So how many arcminutes a pixel corresponds to depends on the resolution of the picture In our low resolution picture one pixel covered 2 arcminutes.

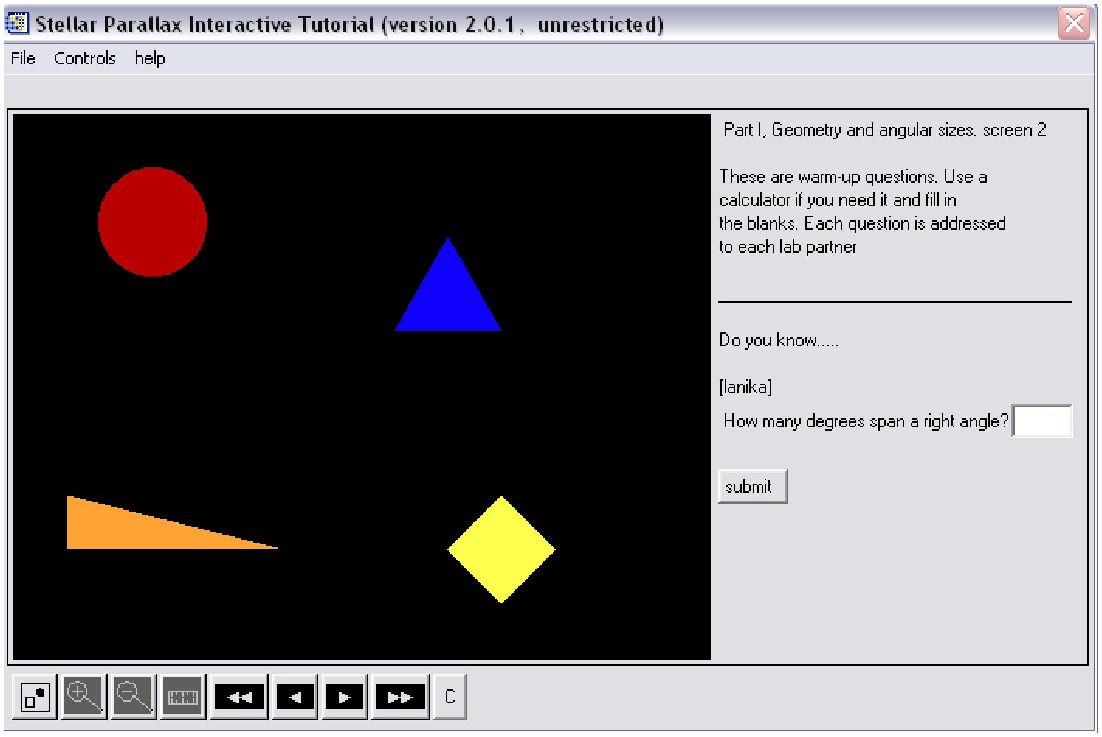
15 pixels = 30'

120 pixels = 30'

And you get better resolution by using a telescope so you can zoom in on a smaller region of the sky In this photo the moon is 15 x 8 = 120 pixels across And every pixel corresponds to 1/4 arcminute, 8 times LESS than the 2 arcminutes per pixel we had before.

Figure 3.5: The concluding scene of a two minute long animation that teaches students about resolution.

SPIRUT users have access to a variety of tools to help them in constructing their knowledge: a drawing tool, a zoom tool, a ruler, navigation buttons and a calculator. This section on *SPIRUT* is concluded by showing some snapshots from scenes from part I, II, and II of the tutorial, parts that were not described in detail (Figures 3.6, 3.7, and 3.8)



Part I, Geometry and angular sizes. screen 2

These are warm-up questions. Use a calculator if you need it and fill in the blanks. Each question is addressed to each lab partner

Do you know.....

[lanika]

How many degrees span a right angle?

Figure 3.6: S*PIRUT* at its most basic, geometry level

Part III, Trigonometric parallax. screen 9

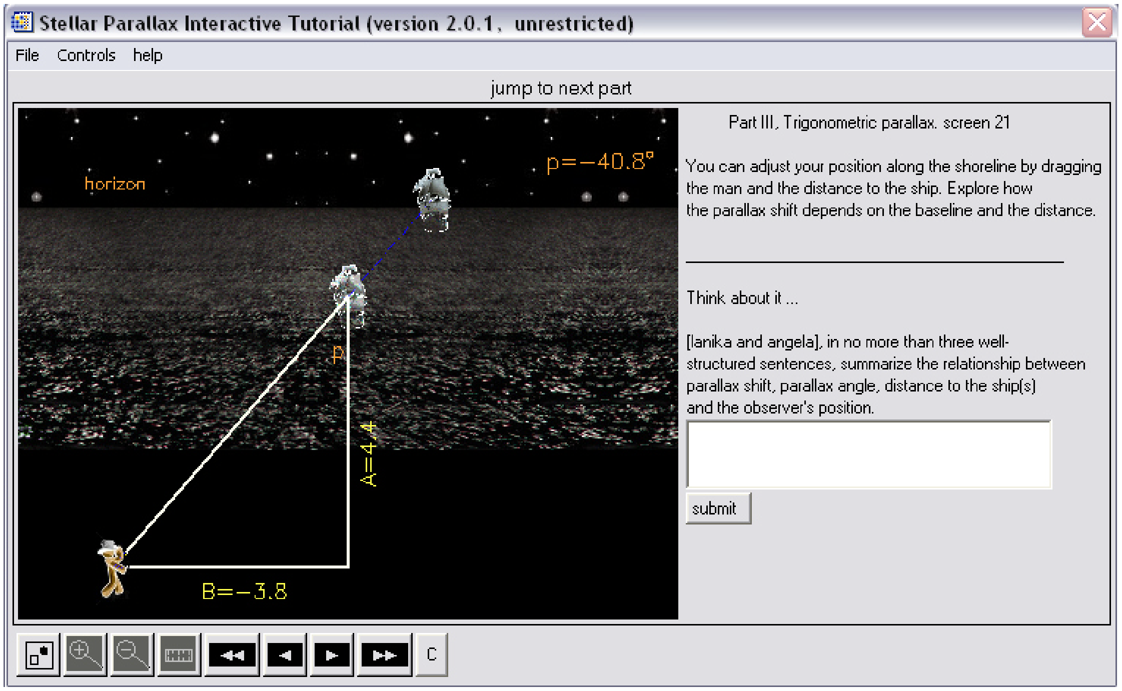
Instructions:
Here is a ship anchored out at sea. You are looking at it from the shoreline. We label your current viewing point as position 2

Select which of the four snapshots would most accurately depict what you would see from position 1. If you need help visualizing, use the drawing pad tool.

Imagine this ...
[lanika], can you visualize how the position of the ship will shift (if at all) with respect to the background stars?

○ ○ ○ ○

Figure 3.7: A scene from part III where more familiar situations are used to link the students' (everyday) experiences to the stellar parallax concept.



Part III, Trigonometric parallax. screen 21

You can adjust your position along the shoreline by dragging the man and the distance to the ship. Explore how the parallax shift depends on the baseline and the distance.

Think about it ...

[lanika and angela], in no more than three well-structured sentences, summarize the relationship between parallax shift, parallax angle, distance to the ship(s) and the observer's position.

Figure 3.8: A scene from a subsequent exercise of part III where the more familiar situations are linked to geometry, and observed angles and shifts.

III. Overview of Data Collection and Data Analysis

Comparing two groups of participants, (1) *Lecture-tutorials* and *SPIRUT*, (2) program-controlled and learner-controlled, (3) peer interaction and individual work, helped to answer the research questions of the study (see Table 3.1 and Figures 3.9, 3.10, and 3.11).

Table 3.1a: *Summary of the research design highlighting utilized methods to answer Research Question 1 of the study.*

<table>
<tr><td colspan="4">Question 1
Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?</td></tr>
<tr><td>Type of Data</td><td>Data Collection</td><td>Data Analysis</td><td>Comparison of the cells and rows</td></tr>
<tr><td>Quantitative</td><td>Pre-tests, post-tests, and retention tests (<i>SPA</i>)</td><td>ANOVA; t-test</td><td>0 X
0 X</td></tr>
<tr><td colspan="4">a. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?
<i>SPA</i> (ANOVA)</td></tr>
<tr><td colspan="4">b. Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?
<i>SPA</i> (ANOVA)</td></tr>
</table>

Table 3.1b: *Summary of the research design highlighting utilized methods to answer Research Question 2 of the study.*

<table>
<tr><td colspan="4">Question 2
Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions in student-controlled or tutorial-controlled SPIRUT?</td></tr>
<tr><td>Type of Data</td><td>Data Collection</td><td>Data Analysis</td><td>Comparison of the cells and rows</td></tr>
<tr><td>Quantitative</td><td>Pre-tests, post-tests, and retention tests (<i>SPA</i>)</td><td>ANOVA</td><td>0 0
X X</td></tr>
<tr><td>Qualitative</td><td>Semi-structured interviews (pairs and individuals)</td><td>Audio recorded, semi-structured</td><td></td></tr>
<tr><td colspan="4">a. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions in student-controlled or tutorial-controlled SPIRUT?
<i>SPA</i> (ANOVA)</td></tr>
<tr><td colspan="4">b. Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions in student-controlled or tutorial-controlled environments of SPIRUT?
<i>SPA</i> (ANOVA)</td></tr>
</table>

Table 3.1c: *Summary of the research design highlighting utilized methods to answer Research Question 3 of the study.*

Question 3 Do students construct greater knowledge of the stellar parallax concept when their learning processes are supported by working with a partner or working independently?			
Type of Data	**Data Collection**	**Data Analysis**	**Comparison of the cells and rows**
Quantitative	Pre-tests, post-tests, and retention tests (*SPA*)	ANOVA	0 0 0 X
Qualitative	Semi-structured interviews (pairs and individuals)	Audio recorded, semi-structured	
a. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are supported by working with a partner or working independently? ***SPA* (ANOVA)**			
b. Do students learn how to reason their answers and transfer knowledge better when their learning processes are supported by working with a partner or working independently? ***SPA* (ANOVA)**			

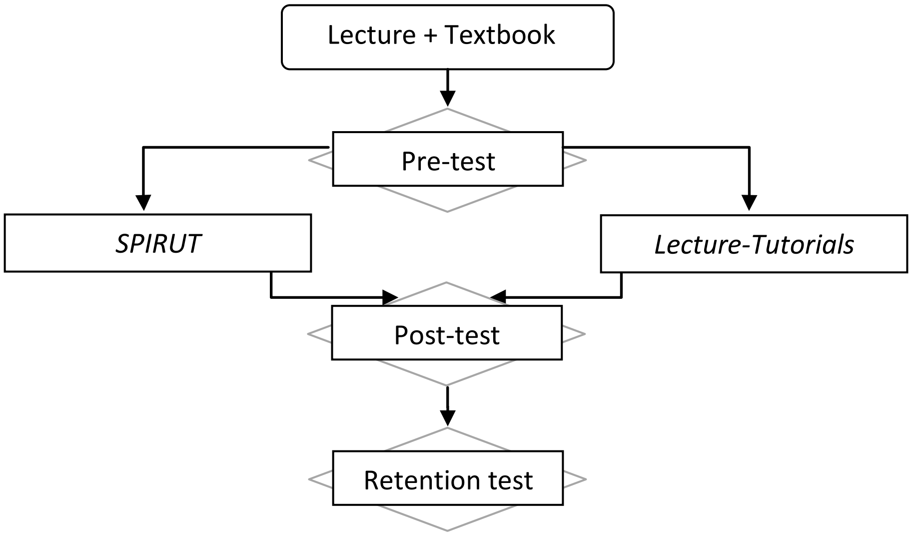


Figure 3.9: Flowchart of the experimental design study answering Research Question 1

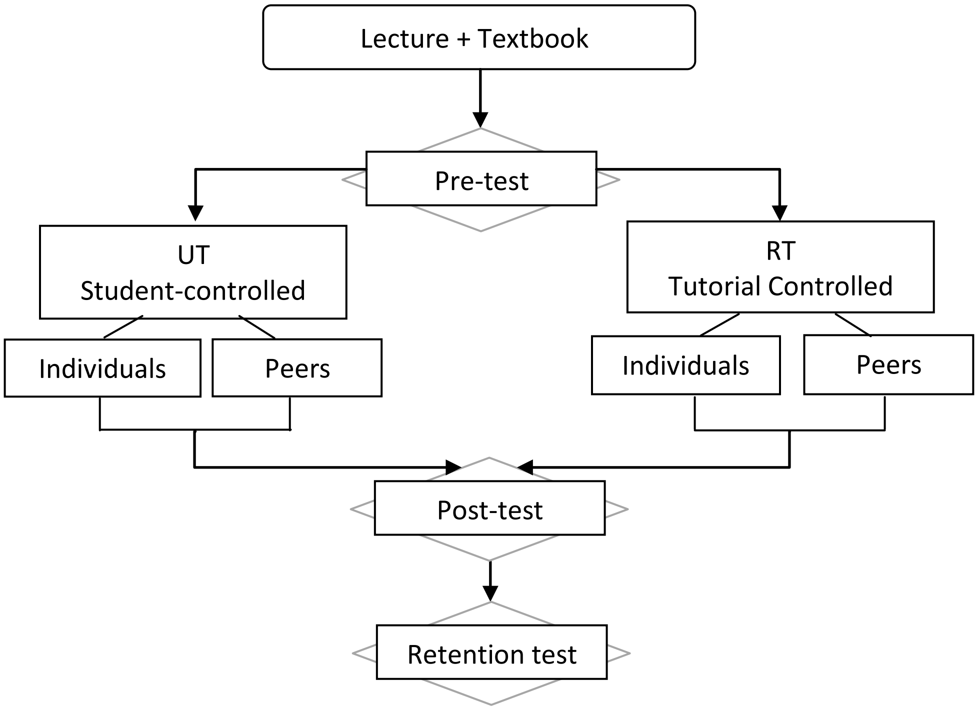


Figure 3.10: Flowchart of the experimental design study answering Research Question 2

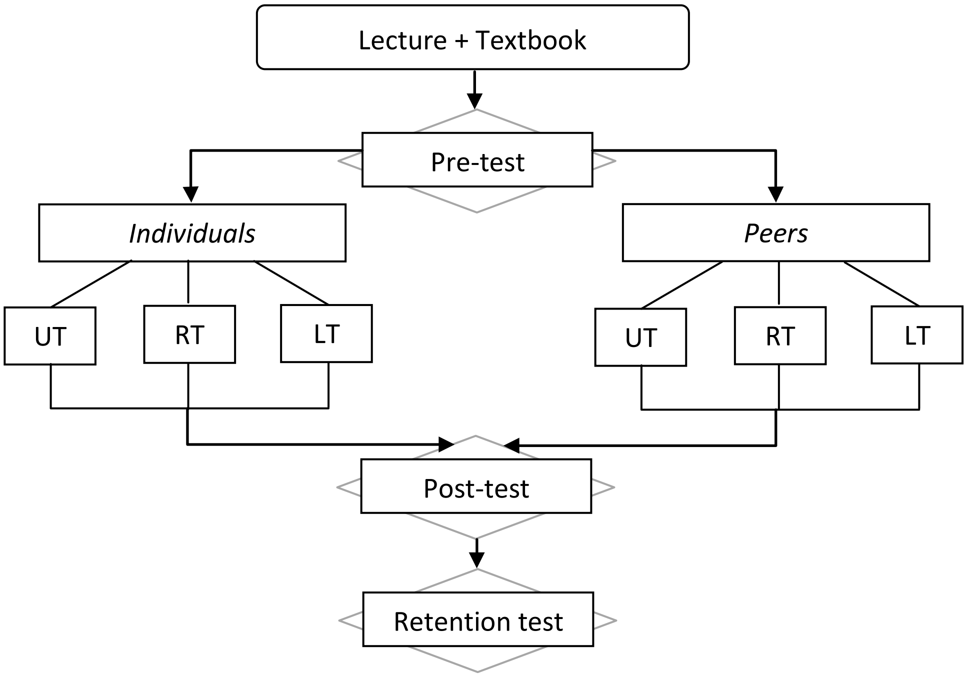


Figure 3.11: Flowchart of the experimental design study answering Research Question 3. *UT* – learner-controlled, *RT* – program-controlled, *LT* – Lecture-tutorials.

IV. Assessment Instrument: Stellar Parallax Assessment (SPA)

The inventory was created for measuring students' conceptual understanding and abstract thinking, or their ability of applying mathematical solutions, for stellar parallax. Currently, there are several inventories and standardized tests that are used to evaluate student achievements in introductory astronomy courses, such as the *Astronomy Diagnostic Test* (*ADT*) (Hufnagel, et al., 1999), diagnostic assessment *Ordered Multiple-Choice* (*OMC*) (Briggs, Alonzo, Schwab, & Wilson, 2006), *Project STAR* Astronomy Concept Inventory (Sadler, 1998), the *Stellar Properties Concept Inventory* (Bailey, 2004), and a *Library of Concepts* (Green, 2003). However, there are no standardized tests or inventories to directly evaluate students' basic understanding of the specific concept – stellar parallax. This first step in finding stellar properties has been almost entirely overlooked by educational research in astronomy. For example, the *ADT* (Hufnagel et al., 1999) consists of 21 questions testing students on a number of aspects in astronomy, and only one question tests students' perception of a shift in a star position based on an observer's change of location. The concept inventories mentioned above test students' understanding of causes of day and night, moon phases, tides, seasons, motion of the celestial sphere, distances inside and outside of the solar system, stars' temperatures, luminosities and magnitudes, gravity, and nature of light (Bailey, 2004; Briggs et al., 2006; Hufnagel et al., 1999; Sadler, 1998). However, there is no inventory that can be used to diagnose students' understandings and misunderstandings of stellar parallax which is a more basic concept that is a basis for building student understanding of more difficult concepts.

At the same time, a number of astronomy education researchers point out the importance of students' understanding of basic scientific concepts (Hufnagel et al., 1999; Green, 2003; Sadler, 1998). Understanding stellar parallax is one of these concepts.

Since there are no existing instruments to measure students' achievements in this area of astronomy courses, the *Stellar Parallax Assessment* was created for use in this study. The *SPI* consists of 15 multiple-choice and one open-ended items based on a combination of the revised questions and exercises in the Parallax tutorial from the *Lecture-Tutorials for Astronomy* (Prather et al., 2006), *Astronomy Diagnostic Test* (*ADT*) (Hufnagel et al., 1999), the *Concept Tests Library* (Green, 2003), and *Well-Structured Problem Solving Process Inventory* (*WPSPI*) (Shin, Jonassen, & McGee, 2003).

All four instruments were validated by their authors and tested for reliability. The *Lecture Tutorials* and *Concept Tests Library* are used in colleges as learning materials, *ADT* was tested with over 1500 students from 17 colleges across the country (Hufnagel et al., 1999) and *WPSPI* was validated by astronomy experts and teachers and tested with 118 students (Shin, Jonassen, & McGee, 2003).

To validate the *SPA,* the instrument was given to six professional astronomers four of whom are professors at universities and teach introductory and upper level astronomy courses and two are college introductory astronomy instructors. Based on the experts' comments, the instrument was revised and pilot-tested on a group of college students. The 22 students enrolled in an introductory astronomy course in a local Midwestern community college were given the version one of the *SPA* (Appendix C).

V. Taxonomic Level of Questions

In his review of literature on effectiveness of computer technology in the social studies Michael Berson (1996) pointed out that while there was evidence of positive learning outcomes as the result of using technology-supported tutorials, “the studies fail to address questions regarding the taxonomic level of questions presented” (p. 489).

Since this study was designed to investigate how students learning processes could be aided by computer-based technologies and peer interaction, we were interested not only in students’ overall performance on the assessments but were eager to investigate what type of learning processes are being engaged and what type of knowledge, factual or abstract thinking, were supported and constructed as a result of the students interactions with the tutorials and with peers.

Thus, 13 questions of *SPA* were further evaluated and subdivided by five professors of astronomy and physics from four different universities and colleges into two groups of subset questions: *Type I* and *Type II*.

Subset questions *Type I* included eight questions that assessed learners’ level of basic mathematical skills and their understanding of basic components of the concept. The *Type I* questions are numbers 1, 2, 3, 4, 5, 6, 8, and 12 on the *SPA* instrument (see Appendix C).

Subset questions *Type II* included five questions that are knowledge transfer questions that require students to think “outside of the box,” require imaginative and abstract thinking to assess the presented situation and be able to apply the stellar parallax concept in different situations others than in which the concept was learned. The *Type II* questions are 7, 9, 11, 13, and 14 on the *SPA* instrument.

In the process of dividing questions into groups, each of the five professors was asked to answer the questions and justify their answers. If a professor felt that the question was a simple recall question and did not require envisioning of a situation, finding an answer through making a conclusion after connecting several pieces of information together, or require a mental application of the stellar parallax method to other situations, these questions were assigned to the *Type I* subset questions group and considered to be a lower-order procedural and factual recall question (Bloom, 1956; Collins, 1985; Graesser & Person, 1994; Gronlund, 1982; Rus, Cai, & Graesser, 2007). Questions that could not be answered by just performing mathematical calculations or by recalling information provided in the textbook or tutorial exercises were considered higher-order questions that involved deeper thinking processes of interpretation and reasoning (Bloom, 1956; Collins, 1985). These questions required to identify and analyze facts, to draw on known information in order to apply it to an unknown situation, to suggest a solution and verify it computationally, and to provide a well-reasoned justification to the given answer. Thus, these questions required students to use their abstract thinking and transfer their knowledge of the concept. Therefore, the questions were assigned to *Type II* subset questions group.

Questions 4, 7, and 12 were not unanimously assigned in one of the two groups of questions. The professors were asked to look at the questions again and comment on how sure they were about their first decision. After this iteration, four out of five professors agreed on the same groups for these questions.

VI. Experimental Design

The experimental comparison design of the study includes quantitative data analysis and qualitative methods for investigation of the research questions.

This mixed methods approach is especially valuable for this study, which was set to examine the effects of technology-supported inquiry-based learning conveyed by *SPIRUT* and peer interaction as well as discover new factors influencing students' academic achievements in astronomy classes.

To answer the research questions and test the set hypotheses stated in Chapter 1, the data was collected using the developed for the study *Stellar Parallax Assessment (SPA)* instrument (see Appendix C). The research design utilizes a classical approach of collecting data during pre- and post- tests in addition to which a third assessment test was conducted to obtain additional information on the participants' retention of knowledge.

The collected statistical data was analyzed using repeated-measures and factorial 2 x 3 ANOVA statistical models complimented by the paired-samples *t*-test. The qualitative inquiry of the study aided in the discovery of emerging themes on students' engagement in their inquiry processes and in peer interaction. In addition, the qualitative inquiry was triangulated with the quantitative tests and provided additional support or refutation to the hypotheses. The qualitative inquiry of the study was based on semi-structured interviews. Participants audio-recorded responses transcribed and scrutinized using combined multiple techniques of pawing, code-a-text, and cutting and sorting. The findings were expected to confirm or reject the predictions set for the investigation.

VII. Participants

Participants of the study were college-level non-science major students from various departments at the University of Missouri. They are Freshmen, Sophomores, Juniors, and Seniors; female and male undergraduate students whose typical age is from 19 to 23 and who are currently enrolled in the *Introductory Astronomy, 1010* course to fulfill the University's general education requirement – completing one biological, physical, or math sciences course.

VIII. Research Procedures

The study was conducted during two days on September 9th and 10th, 2011. The study was conducted in a number of steps:

1. All participants received a 20-minute lecture on the stellar parallax concept
2. All students immediately received a two-page printout from the course textbook explaining the stellar parallax concept
3. All students were given a pre-test which they answered within 10-15 minutes
4. Within next two days all students were administered treatments:
    a. Group 1: *Lecture-Tutorials,* a paper-based exercises, a control group
    b. Group 2: *SPIRUT,* a computer-based learner-controlled environment
    c. Group 3: *SPIRUT,* a computer-based program-controlled environment

       NOTE: each of the three groups were also sub-divided into pairs and Individuals groups

5. After completing of each treatment, the students were administered the post-test
6. Eight weeks later during a regular class the students were given the retention test
7. A week later, thirty six out of 199 students were selected and participated in follow-up semi-structured interviews.

During Step 1 students listen to an instructor's presentation and explanation of the stellar parallax using seven PowerPoint slides. Students were allowed to ask questions if they need any clarifications.

During Step 2 step of the procedure the participants were asked to read two pages from the course textbook chapter that explained the trigonometric and stellar parallax concepts with mathematical derivations of the formula (see Appendix A). This step took less than 10 minutes.

During Step 3 students were given the *SPA* pre-test designed to assess their current level of understanding of the concept (see Appendix C). The pretest was also designed to test students' ability to apply simple mathematical calculations in order to find distances to stars and necessary units' conversion. The pretest was placed between the students' reading of the textbook and working with the tutorials procedures. Such order allowed obtaining of more accurate information on what knowledge students constructed upon the lecture and the reading of the textbook. This step took less than 15 minutes.

During Step 4 within next two days students participated in treatments in which they were randomly assigned. On average students spent about 40 minutes working with *Lecture-Tutorials* (see Appendix B) and *SPIRUT*.

During Step 5 students completed the post-test. It took about 15 minutes for each student to answer all the questions on *SPA*.

During Step 6 students were asked two days in advance to come to class for an important class activity. As a result of it, most of the students completed the retention test within allotted 20-minute time slot during a lecture. Originally 212 students participated in the study, however not all of them submitted the retention test. As a result, their previous records were removed and only 199 students' assessment tests were used for the data analysis.

During Step 7, thirty-six students were selected for the semi-structured interviews based on their negative or low, medium or high gain. The interviews took place in the University observatory and lasted between 15 and 25 minutes (see Appendix D). A schematic representation of the study's procedures and the number of participants in each step of the data collection and data analysis are shown in Figure 3.12.

As a result of a short of period of time passing between administration of the pre- and post- tests, there is a possible threat of testing effect to the internal validity of the study. However, use of second version of the *SPA* minimizes the affect. The second version of the inventory uses the same conceptual questions with adjusted mathematical parameters, dates of the data collection, and names of the stars. The developed instrument allows evaluating students' understanding of the concept as well as the students' ability to apply the concept for mathematical solutions.

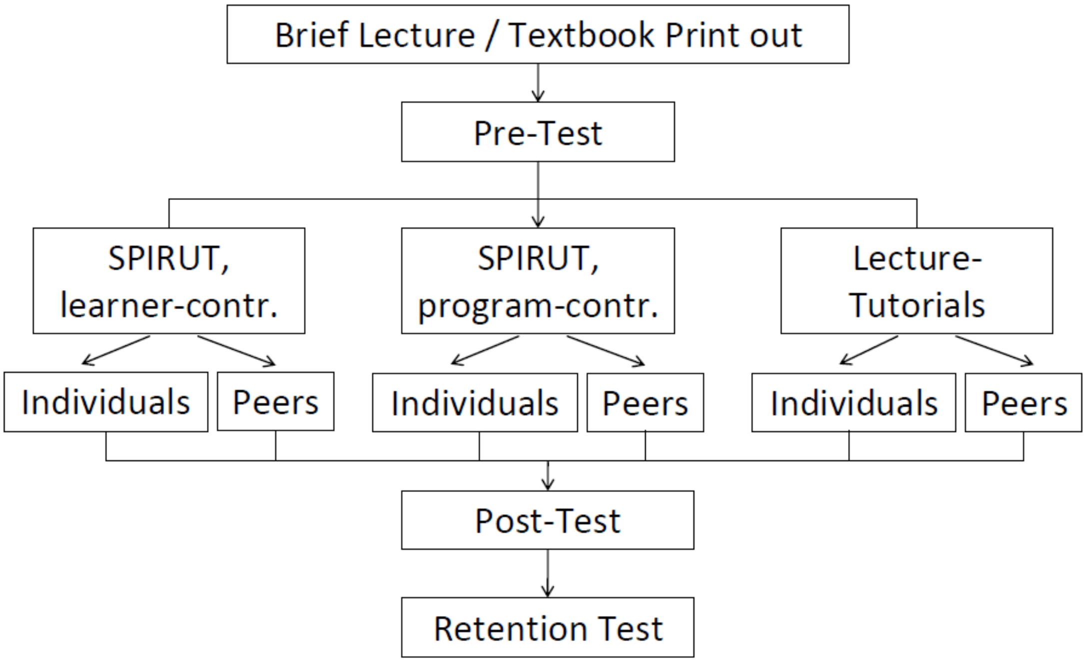


Figure 3.12: Schema of the study's procedures with sample sizes for each unit of data collection and data analysis.

IX. Quantitative Data Analysis

IX.I. 2x3 Repeated-Measures Factorial ANOVA

The quantitative analysis of the research follows a 2 x 3 factorial experimental design also often referred to as repeated-measures ANOVA model. Repeated-measures ANOVA is used when two or more independent variables or factors are present in the experiment and the same unchanged variable is being measured repeatedly. This design is one of many types of factorial designs where each of several independent variables is repeatedly measured using the same participants divided into groups (Field, 2005) and thus measures variability between and within the groups.

Moreover, the repeated-measures design is also an effective method for examining interaction effects among the variables (Trochim, 2000). Since reality is not based on a random occurrence of a single factor influencing the next upcoming factor, but is rather a constant and simultaneous multi-variable interaction of influencing factors it makes sense to examine the variables as they interact. In this case, "factorial designs are closer approximations to reality" (Rutherford, 2001, p.43). Consequently, Keppel (2004) highlights the importance of interpreting effects of each independent variable in concert with another independent variable so the influence of the interacting independent variables on the dependent variable can be measured. In other words, the uniqueness of factorial designs is in the ability to examine how combinations of manipulated factors - independent variables - influence behavior dependent variables (Rutherford, et al.).

Interaction among predictors and the ability of factorial designs to contrast means of a number of experimental conditions without increasing Type 1 error - observing a difference when there is none - makes the design a popular model of statistical analysis (Rutherford, et al.). Factorial analysis of variance is commonly used in social sciences in the field of education and technology.

ANOVA statistical data analysis models are used in a range of fields passed way outside of social sciences. For example, ANOVA data analysis proved being a successful model in studies on pilots tutoring techniques and in a comparative study of experts versus novice pilots.

In first case, factorial ANOVA was used by Kasarskis, Stehwien, Hickx, and Aretz (2001) in their experimental study of influence of pilot experience levels (novices vs. experts) on quality of their aircrafts' landing performance (good vs. poor) using a simulated environment found not only differences and correlations between the groups and their performances but also discovered an interaction among particular items on the test. This interaction helped the researchers explain why experts performed better compared to novice pilots.

In the second case, a study conducted by Purcell and Andre (2001) used a factorial ANOVA design to compare visual and auditory callouts used by commercial pilots. While there was no significant difference found in main effects, the means comparison revealed a pattern where certain details displayed higher numbers for the callout conditions than the no callout, as well as specific visual and audio callouts showed differences when compared to the baseline of no callouts (Purcell & Andre, et al.).

While ANOVA models are robust tests, these models compare different groups for the same test and compare the same group for different tests. Our study required an additional comparison of different groups for differences in tests is what we called gain. This was achieved by analyzing the data using the paired-samples t-test statistical model.

IX.II. ANOVA: Data Analysis Procedures

The statistical ANOVA model examined main effects – the difference of marginal means between groups in columns and rows, of two independent variables with two conditions to two quantitative dependent variables.

Independent Variables (IVs):

1. Computer Interaction Factor
    - (a) *SPIRUT*, (b) *LTs*

2. Computer Factor (*SPIRUT*)
    - (a) Learner-control, (b) Program-control

3. Social Interaction Factor
    - (a) Individual, (b) Peer Interaction

Dependent Variables (DVs):

1. *Type I* questions, simple operations
2. *Type II* questions, knowledge transfer of the stellar parallax concept

In order to test the influence of the two independent variables on each of the dependent variables, the model was run twice separately for each of the dependent variables.

The investigation was based on both equal and unequal sample size groups of students. Total of 199 participants were separated in the following groups:

- *SPIRUT:* 135 students
- Program-controlled: 67 students
- Learner-controlled: 68 students
- LTs: 64 students
- Individuals: 99 students
- Pairs: 50 pairs (100 students)

To estimate the internal validity of the study, statistical values for factorial ANOVA in this study was obtained using *GPower* software, which predicted the following statistical values for a total sample size of 120: effect size - 0.31; alpha level - 0.05; power - 0.8097; and critical *F* (3,116) = 2.6828. This output satisfies the internal validity of the study.

IX.III. Paired-samples *t*-Test

The paired-samples *t*-test is an appropriate model to employ for analysis of a single predictor in pared observations (O'Rourke, Hatcher, Stepanski, 2005). In the case of the study, the model was used to analyze differences in tests of the same participants over time on three assessment tests. Scores achieved on each assessment test by every individual participant were correlated with the same participant previously obtained scores. The use of the *t*-test in the study allowed the researchers to conduct deeper and more meaningful analysis of the collected data.

X. Qualitative Analysis

While statistical data can reveal interactions and levels of significance of change in students' learning outcomes, it does not provide insight into the learners' ways of thinking and working as well as levels of engagement with the assignment. Qualitative inquiry can discover new themes and patterns, categories and ideas, and "in general uncovers better understanding of a phenomenon or process" (Newton Suter, 2006, p.327).

X.I. Stratified Random Sampling

212 students enrolled in the course were not equally distributed among the freshmen, sophomore, junior, and senior classes. Each group is not equally represented: there are 50 freshmen, 20 sophomores, 10 juniors, and 40 seniors. To assure equal variability among the groups during the experiment, the stratified random sampling or randomized block design (Trochim, 2000) is used to divide the class into four homogeneous subgroups of 30 participants in each. This type of sampling allows for equal distribution of all key subgroups of the population (Trochim, et al.) thus enabling investigation of the treatment variability within a particular group of students.

The 212 enrolled students are divided into four groups as follows:

$$N_{120} = N^1_{30} + N^2_{30} + N^3_{30} + N^4_{30}.$$

A simple random sample is applied to ensure equal representation of each demographic group within each strata:

$$f = n/N,$$

Where f is a number of subgroup representatives in strata, n is a sample size, N is a population size.

The students' population is sampled as follows:

Freshmen – 50/120 = 42%, within the strata: f_{13}=30/100*42;

Sophomores – 20/120 = 17%, within the strata: f_5=30/100*17;

Juniors – 10/120 = 8%, within the strata: f_2=30/100*8;

Seniors – 40/120 = 33%, within the strata: f_{10}=30/100*33.

As a result of the small sample of sophomore and junior demographic subgroups, variability of the treatment within these subgroups is negligible and, therefore, statistically invalid. These subgroups, however, can be assigned to the two major subgroups - freshmen or seniors respectively. Thus, the study investigates the treatment variability between two major subgroups of students of first, second, third and fourth years of the school enrollment.

X.II. Systematic Random Sampling

To assure randomization within each group, the systematic random sampling technique is applied within each subgroup. To randomly select students from demographic subgroups to each of the experiment groups, a number of the subjects in the population (*N*), sample size in a treatment group (*n*), interval size, *N*/*n*, (*k*), and a random integral (*i*) arbitrarily chosen between 1 and *k* (Trochim, 2000) are used. Every k^{th} unit is selected for the sampling from the alphabetically and numerically ordered class roster arranged by students' years in college, (Table 3.2).

Freshmen: $N = 50$, $n = 13$, $k=4$, $i = 3$; thus the selected sampling units are 3, 7, 11, 15, and so on.

Table 3.2: *Systematic random sample of freshmen.*

1	**11**	21	**31**	41
2	12	22	32	42
3	13	**23**	33	**43**
4	14	24	34	44
5	**15**	25	**35**	45
6	16	26	36	46
7	17	**27**	37	**47**
8	18	28	38	48
9	**19**	29	**39**	49
10	20	30	40	50

Consequently, other groups of students will be assigned to the subgroups as follows:

Sophomore: N = 20, n =5, k=4, i=3;

Juniors: N = 10, n =2, k=5, i=4;

Seniors: N = 40, n =10, k=4, i=3.

This easy and precise technique assures random sampling of all subjects within each subgroup (Trochim, et al.).

X.III. Follow-up Interviews

The follow-up interviews were used to collect attitudinal and perceptual data (Slavin, 1984). The interviews were intended to introduce additional richness and completeness to the observed data (Slavin, et al.). The follow-up interviews added more information on students' experience working with and learning from the *SPIRUT* and *LTs* as well as working with a peer or independently. Students' answers helped the investigators gain insight on why the students achieved higher or lower results on the post-test and/or retention test. A pattern noted in the students' answers allowed combining the answers to reveal an emerging theme. The found themes explained the students' interactions with the tutorials, collaboration with peers, and difficulties or successful approach with transforming visual representations into mathematical solutions. More details about the asked questions are described in the *Instruments* section in this chapter.

Based on the students' performances during the treatment, students who gained the most and the least were selected for the follow up interviews. The semi-structured follow-up interviews helped to establish why some students benefited the most while some students fared poorest (Appendix D). Students asked to come for the interviews will be offered astronomy posters as incentives. If students would like to participate in the interviews and receive the posters but are precluded from participating in the interviews, they will be offered an opportunity to write a short essay where they can answer the same questions that are designed for the semi-structured follow-up interviews. The process of normalized gain was used to determine the interviewees.

The study participants were divided in three groups according to their knowledge gain: high, medium, and low, based on the results of the pre- and post tests. Zelick, Schau, and Mattern (1999) used the normalized gain approaching to evaluate learners' conceptual gain in their study. The normalized gain index is found by setting up a proportion where actual average students gain is compared to the maximum possible gain:

$$g = (post\%-pre\%)/(100-pre\%),$$

where g is a gain index with 0 is no gain and 1 is a maximum gain.

In this study the normalized gain will be considered high if it falls between *1* and *.67*, medium is *.66* to *.34*, and low gain corresponds to *.33* to *0*.

Thus, 36 selected students who showed high (12 students), medium (12 students), and low (12 students) gains were interviewed eight weeks after the testing. The length of each interview with 20-23 semi-structured and one-two open-ended questions lasted about 20-25 minutes. All interviews were audio recorded and transcribed.

Out of the 36 students, 18 students were selected from those who worked individually and 18 - from the peer interaction groups. The participants from each peer interaction pair were selected based on their individual normalized gains regardless of their peer-partners achievements. The follow-up interviews helped the investigators to understand why the gain did or did not occur.

The interviews helped to establish what aspects of the tutorials are beneficial to the learners' experiences and what factors preclude the students from obtaining better results. Since the qualitative inquiry of the study aimed to discover themes and motifs, the medium gain was noted but did not represent most interest in this study.

Respondents reports gathered from the follow-up interviews were especially helpful for finding answers to the third questions of the study: whether peer interaction has positive contribution to the learners' construction of knowledge.

X.IV. Qualitative Inquiry Analysis

The transcribed interviews were coded by two co-raters in four steps. During a first meeting one randomly selected interview was analyzed on student's responses to the 34 questions of the semi-structured interview that were repeatedly asked in every interview. The student's answers on these questions were highlighted in the printed transcript, coded and entered in an excel spread sheet. Additionally, provided information and remarks by the student were noted as a theme. These "themes," were also color-coded and entered in a separate spreadsheet designated for emerging themes. Each student’s answer was negotiated between the two co-raters until they achieved consensus about how to code the replies.

In the second step of the qualitative analysis, after identifying themes and coding three interviews, co-raters came to agreement about which categories should be included in a created table so it would be consistent throughout all interviews.

In the third step, after coding was finished, the codings on the two co-raters spreadsheets were compared and several discrepancies were discovered. The inter-rater agreement of the qualitative analysis was 86% at that time.

Thus, in the fourth step every discrepancy was reviewed and new codes negotiated and until co-raters came to an agreement. Thus final inter-rater agreement of the study was 100%. However, while co-raters agreed on most of the codes, thirteen out of 1,224 codes were classified as “conditional agreement.” In these cases, six discrepancies in coding were based on students’ uncertainty and unclear expression whether they would rather work individually or with a partner, three - whether students preferred learner-controlled or computer-controlled tutorial, and four - whether students preferred answering multiple-choice or essay questions.

After the coding was finalized, emerging themes were discussed.

XI. Summary

The study was designed to examine how a computer-based interactive tutorial can aid students learning processes and facilitate their engagement to abstract thinking during the transfer of knowledge, helping to construct procedural knowledge of the scientific construct, and practice their mathematical skills. In addition, the study aimed to investigate peer interaction and its effect on students learning.

The study involved the analysis of two independent variables (predictors): (1) students' work with *SPIRUT* as compared to work with *LTs*; (2) students work with two versions of *SPIRUT*, a learner-controlled and a program-controlled; (3) and students' social peer interaction in pair groups as compared to other students' individual work on the task. Students' learning outcomes were evaluated based on their achievements as they were measured by *Stellar Parallax Assessment* (*SPA*).

With this in mind, we designed the study to achieve internal and external validity for statistical tests and qualitative inquiry. The collected statistical data was analyzed using repeated-measures and factorial 2 x 3 ANOVA models to find differences in means of participants' performance on three assessment tests within their assigned groups and in between the groups of participants. An independent-samples t-test was performed to compare these different groups for statistical differences in tests (their gains). The statistical tests also revealed whether there was an interaction and dependency among the factors.

Direct observations and follow-up interviews were included in the research design to find emerging themes of participation and engagement, and to identify patterns of peer interaction and individual work with *SPIRUT* and *Lecture-tutorials*.

The utilized research methods allowed the investigators to recognize learners' motivations, attitudes, and their experiences based on their social and computer-based interactions.

In compliance with the University of Missouri campus' institutional review board (IRB), a treatment of human subjects proposal was developed and approved by the review board (IRB # 1135257). No data collection was undertaken until approval had been granted and no students participated in the investigation without their written voluntary consent.

CHAPTER 4

DATA ANALYSIS AND RESULTS

I. General Findings and Chapter Structure

In this chapter, the results of statistical analysis are presented, as well as findings from the qualitative investigation part of the study. The data analysis is used to answer the following questions in the study:

1.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

  1.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

  1.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

2.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions in learner-controlled or program-controlled *SPIRUT*?

2.2 Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions in learner-controlled or program-controlled *SPIRUT*?

2.3 Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions in learner-controlled or program-controlled environments of *SPIRUT*?

3.1 Do students construct greater knowledge of the stellar parallax concept when their learning processes are supported by working with a partner or working independently?

3.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are supported by working with a partner or working independently?

3.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are supported by working with a partner or working independently?

The main finding of administering the treatments and analyzing the data was that both the computer based and the paper based treatments increased student knowledge over what they had learned in traditional class sessions. Both groups showed significant improvement in conceptual understanding of the stellar parallax concept on top of the knowledge they gained during a 20 minute class lecture. ($M_{\text{pre-test}} = 6.726$, $M_{\text{post-test}} = 7.834$, $p = .000$, $p < .001$)

In addition, the group that used the computer-based tutorial gained more than the paper-based group (research question 1). There was no difference in performance amongst the students in the computer-based group depending on whether the tutorial was run in the restricted, or unrestricted (student controlled) mode of operation (research question 2).

It was found that individuals outperformed students working with peers (research question 3). The differences were not significant at the 95% level though. However, since the finding is counter-intuitive these results have been made the focus for the qualitative part of the study.

For the three main research questions a repeated-measures ANOVA was used to investigate whether there was a main effect and to investigate whether the gains within groups were significant. Six subset questions were answered using a 2x3 factorial ANOVA. To compare the significance level of the difference in gains between the groups, the paired-samples *t*-test was used. For the research subset questions, (questions 1.2, 1.3, 2.2, 2.3, 3.2 and 3.3) the 13 assessment questions were subdivided into two groups (*Type I* and *Type II*) according to type of knowledge required to answer them as specified in Chapter 3.

This chapter starts with a brief description on the testing to establish the validity of the underlying assumptions for statistical analysis. Next the results of the research questions are presented. And lastly the data of the qualitative study is presented. These quantitative and qualitative results are discussed and interpreted in Chapter 5.

II. Statistical Analysis

II.I. Validity of Assumptions

In order to answer the three main and six subset questions of the study, two types of statistical approaches and a qualitative data analysis were used. To identify how students performed on three assessment tests over a six-week period at each point of time, ‘repeated-measures ANOVA’ was used. For each of the three main questions the analyzed data was scrutinized to test the validity of four assumptions: multivariate normality, independence of observations, sphericity (homogeneity of covariance), and between-groups equality of variance.

Distribution of scores within and between groups was scrutinized using graphical methods, histograms and normality plots, which displayed a normal distribution of the scores. Kurtosis and skewness of scores were also examined and the variables were found reasonably close to normal falling within the range of -2 to +2. In addition, Kolmogorov-Smirnov test of normality was performed which revealed the actual distribution of the variables were equal to the expected distribution ($p = .200$, $p > .05$) indicating that the multivariate normality assumption was not violated.

All participants in the study belonged to one population: students who chose to take an introductory astronomy course. These students were randomly and independently from their prior knowledge of the concept assigned to a total of six different groups, thereby minimizing the potential violation of the independence of observations assumption.

Mauchly’s test of sphericity was performed to look for a potential common violation of the homogeneity of covariance assumption. The assumption of homogeneity of covariance has not been violated on any of the three questions, as indicated by not-significant values for the approximate Chi-Square values and their associate significance values on Mauchly’s test.

Examined standard deviations for each group using Levene’s test indicated the equality of groups in most of the cases (with a possible exception discussed at research question 1.2) and therefore, the assumption of between-groups equality of variance was not violated.

II.I.I. Research Question 1

1.1 Computer-based and Paper-based Tutorials

Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

It was found that the students using the computer-based tutorial significantly ($p = 0.034$) outperformed their peers using the paper-based tutorial on assessments immediately following the treatment. However, the gains were found to be short-lived for both groups.

Table 4.1: *At a glance: significance levels for Research Question 1.1 pertinent to type of treatment*

Type of Treatment				
Main effect: assessments	**Main effect: treatment**	**Gain *SPIRUT* > 0**	**Gain *LTs* > 0**	**Gain *SPIRUT* > Gain *LT***
$p < .001$	$p = .240$	$p < .001$	$p =. 004$	$p = .034$

In order to answer research question 1.1, 2 (groups) x 3 (assessment tests) ANOVA with repeated measures on three assessment tests were performed on students' achievements as measured by the students' scores on the *Stellar Parallax Assessment* test (Appendix C). To investigate further whether differences in means between groups were significant, the paired-samples *t*-test was performed on two treatments groups, (Table 4.1).

Was there a main effect?

The multivariate statistics test using Wilks' Lambda test indicated a significant main effect of assessments test ($F_{Wilks'\ Lambda}$ (2, 196) = 52.664, $p = .000$, $p < .0005$), where the effect of differences between assessment tests accounted for 35 percent of variance ($\eta_p^2 = .350$). Students received significantly different scores on the assessment tests as is indicated by statistically significant means (Table 4.2, Multivariate test).

Table 4.2: *Multivariate test: Three-way Repeated-Measures ANOVA for SPIRUT and LTs on three assessment test, (n = 135 and 64 per cell respectively).*

Source	Wilks' Lambda	*F*	*Hypothesis df*	*Error df*	*p*	Partial eta^2	Power
test	.650	52.664*	2.000	196.000	.000*	.350	1.000
test by *SPIRUT_LTs*	.976	2.390	2.000	196.000	.094	.024	.479

Note. *$p < .0005$.

In agreement with Willks' test, Greenhouse-Geisser statistics test determined that students performance on three assessment tests were statistically significantly different $F(2, 394) = 51.530$, $p = .000$, ($p < .0005$), which was a close to moderate effect (η_p^2 = .207) with strong statistical power (observed power, 1.000), (Table 4.1, *Within Subjects*).

At the same time, there was no significant main effect between the *SPIRUT* and the *LTs* groups ($F(1, 197) = 1.386$, $p = .240$, $p > .05$), which was a small effect (η_p^2 = .007) that accounted for .7 percent of variance and which did not have sufficient statistical power (observed power, .216), (Table 4.3, *Between Subjects*). This may suggest that students in both groups had constructed comparable knowledge of the stellar parallax concept.

Table 4.3: *Within and Between Groups: Three-way Repeated-Measures ANOVA for SPIRUT and LTs on three assessment test, (n = 135 and 64 per cell respectively).*

Source	Type III *SS*	*df*	*MS*	*F*	*p*	Partial eta^2	Power
Between Subjects							
SPIRUT_LTs	10.022	1	10.022	1.386	.240	.007	.216
Error$_{BS}$	1424.175	197	7.299				
Within Subjects							
test	183.161	1.970	92.972	51.530*	.000	.207	1.000
test by *SPIRUT_LTs*	7.532	1.970	3.823	2.119	.122	.011	.431
Error$_{WS}$	700.221	394	1.777				

Note. *$p < .0005$.

Overall, these two main effects are not dependent upon each other, as indicated by the not significant level on the multivariate test measuring the interaction between the two types of learning experiences given to students as well as by their performance on the assessment tests (*F(*2, 196) = 2.390, *p* =.094, *p* > .05), which was a small effect (η_p^2 = .024), (Table 4.2).

However, small effect sizes and insufficient statistical power displayed by non-interacting variables (two groups and three assessment tests) indicate a potential problem in the research design that is discussed further in Chapter 5. However, lack of interaction of groups across time does not preclude an investigation of means within the groups on three assessment tests.

Were the treatments successful?

Further investigation of the differences between assessment tests within groups using the Bonferroni adjustment post hoc tests revealed that students in both groups statistically significantly improved their scores on the post-test. Students who worked in *SPIRUT* group gained 1.292 points (*p* = .000, *p* < .0005) on the post-test while students from *LTs* group gained .719 points (*p* = .004, *p* < .05). It also should be mentioned that students in both groups statistically significantly lowered their scores on the retention assessment test administered eight weeks later (Table 4.4, *Pairwise Comparison*).

Despite the gains achieved by both groups on the post-test, mean scores of both groups were significantly lower on the retention test when it was compared to the post-test: lower by 1.21 points for the *LTs* group (*p* = .001, *p* < .05) and by 1.61 points for *SPIRUT* group (*p* = .000, *p* < .0005) indicating that neither treatment was successful for retention of students' knowledge of the stellar parallax concept. (Fig. 4.1)

Table 4.4: *Pairwise comparison of two groups: SPIRUT and LTs, (n = 135 and 64 per cell respectively) on three assessment tests*

groups	(I) test	(J) test	Mean Difference (I-J)	Std. Error	*P*	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
SPIRUT	1	2	-1.292*	.152	.000	-1.659	-.925
		3	.321	.168	.175	-.086	.727
	2	1	1.292*	.152	.000	.925	1.659
		3	1.613*	.166	.000	1.212	2.014
	3	1	-.321	.168	.175	-.727	.086
		2	-1.613*	.166	.000	-2.014	-1.212
Lecture-Tutorials (LTs)	1	2	-.719**	.221	.004	-1.252	-.186
		3	.489	.244	.141	-.102	1.079
	2	1	.719**	.221	.004	.186	1.252
		3	1.208*	.241	.000	.626	1.790
	3	1	-.489	.244	.141	-1.079	.102
		2	-1.208*	.241	.000	-1.790	-.626

Note. **p* < .001, ** *p* < .05.

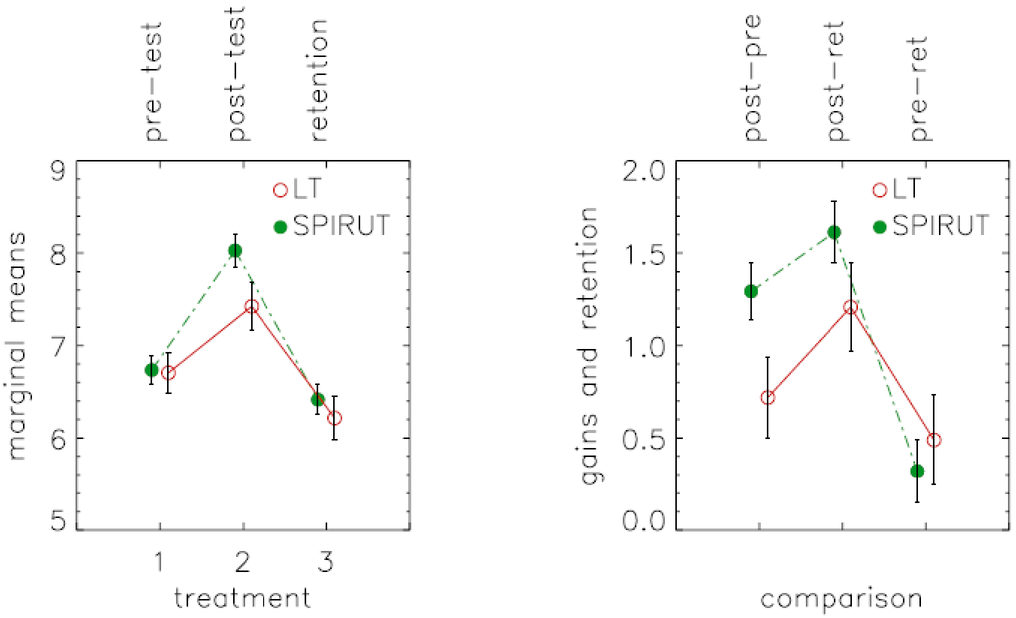


Figure 4.1. Overall students' performance of the two treatment groups (*SPIRUT* and *Lecture-Tutorials*) on the three assessment tests: 1 – pre-test, 2 – post-test, 3 – retention test (left panel). Overall students' gains of the two treatment groups (*SPIRUT* and *Lecture-Tutorials*), (right panel). The error bars in this and similar figures represent the standard error as produced by ANOVA.

Were there differences in gains between the groups?

Distribution of scores revealed by the pairwise comparison test indicated - with 95% confidence - that actual gain on the post-test is within .925 and 1.659 points for the *SPIRUT* group and between .186 and 1.252 for the *LTs* group, respectively (Table 4.4). The small overlap of two distributions shown in figure 4.2 visualizes that one gain is indeed significantly different from the other (Fig. 4.2).

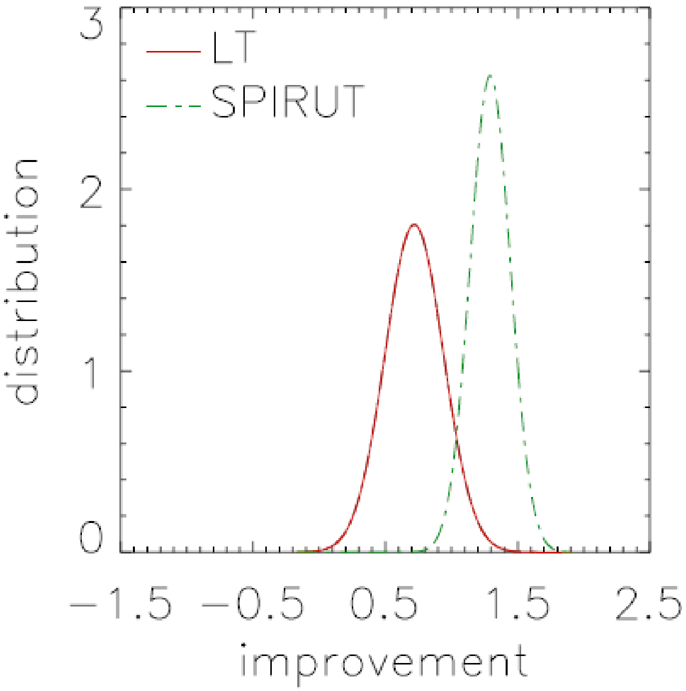


Figure 4.2: The distributions for the gains following treatment for the *SPIRUT* group (right curve) and the *LTs* group (left curve) approximated by Gaussians based on the standard errors given in Table 4.4. The overlap between these approximate curves put the significance level of the difference at 0.034. The horizontal red arrow indicates the 95% confidence interval for the *LTs* gains.

To investigate this effect further, an independent sample *t*-test was performed to compare differences in gains between two groups on the post-test. The statistics test revealed that the differences in gains, as measured by difference in points on achieved scores between post- and pre- tests, were indeed statistically significant between *SPIRUT* (*M* = 1.292, *SD* = 1.714) and *LTs* (*M* = .719, *SD* = 1.873), (*t* (197) = 2.137, *p* = .034), (Table 4.5). These results show that computer-based tutorials helped students gain more knowledge of the stellar parallax concept (Fig.4.3) than when given paper based tutorials.

Table 4.5: *The paired-samples t-test on students' gains in SPIRUT and LTs groups*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
SPIRUT	135	1.292*	1.714	1.002	1.584
Lecture-Tutorials	64	.719*	1.873	.0257	1.181

Note. *$p < .05$, $p = .034$

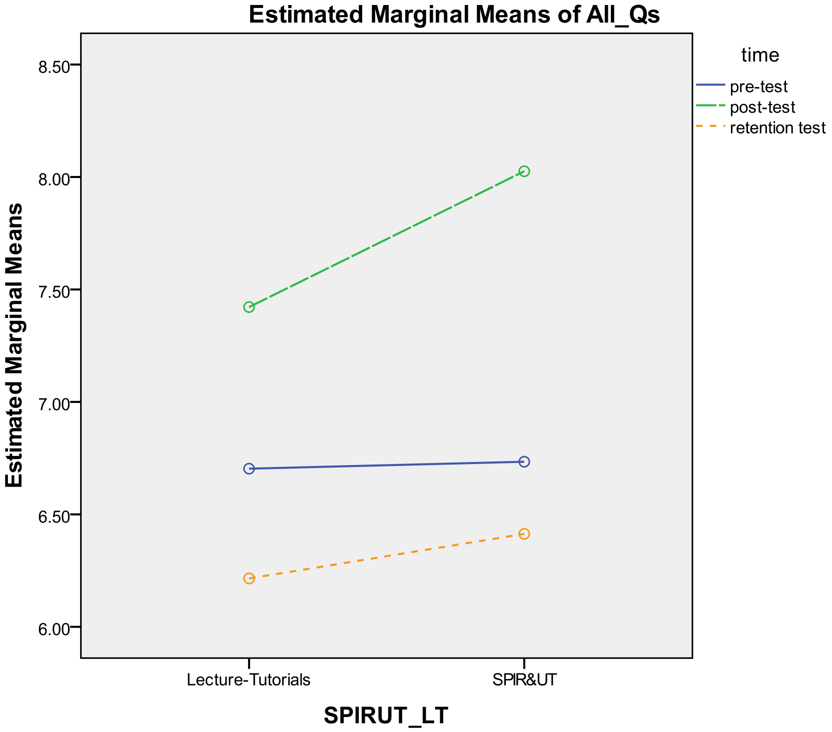


Figure 4.3: The plot was generated in factorial ANOVA to demonstrate means of two groups on three tests. Students in *SPIRUT* group performed considerably better ($p = 0.034$) on the post-test compared to their test results on the pre-test but neither of the groups retained their knowledge after an eight-week period.

Summary

Thus it can be concluded that students indeed increased their knowledge of the concept as a result of their interactions with the tutorials, and that computer-based tutorials were more effective than paper-based tutorials for knowledge gains tested immediately after the treatment. However, students' knowledge of the concept was not long-lasting as indicated by the retention test. Eight weeks later students in both groups received lower scores on the assessment test which demonstrated that neither of the tutorials helped students in constructing a long-lasting understanding of the concept.

1.2 Type I Knowledge Gain

Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

The analysis did not reveal a significant difference in gains between the two treatments. While ANOVA indicated a significant main effect, Bonferroni pairwise comparisons tests showed that the main effect observation likely resulted from partial violation of Levene's assumption, (Table 4.6).

A 2 (groups) x 3 (assessment test) factorial ANOVA was performed on students' learning achievements as measured by the set of questions on the three assessment tests with only *Type I* questions as input (see Chapter 3 for the division into *Type I* and *Type II* questions).

Table 4.6: *At a glance: significance levels for Research Question 1.2 pertinent to type of treatment*

Type of Treatment				
Main effect: assessments	**Main effect: treatment**	**Gain *SPIRUT* > 0**	**Gain *LTs* > 0**	**Gain *SPIRUT* < Gain *LTs***
$p < .001$	$p = .022$	$p < .001$	$p <= 0.020$	$p = .471$

Was there a main effect?

There was a significant main effect of assessment tests on *Type I* questions (*F(*2, 591) = 16.149, p = .000, p < .0005) indicating that students results on the assessment tests were statistically significantly different across time (Fig. 4.4). Furthermore, results showed that there appeared to be a significant main effect of type of tutorials on students learning achievements (*F(*2, 591) = 5.307, p = .022, p < .05) but with a very small η_p^2 value (η_p^2 = .009) (Fig. 4.5). However, the (apparent) main effect of treatment might be artificially bolstered by the fact that the *LTs*-group already did worse on the pretest, restricting their potential gains on the post-test. Given the smaller questionnaire size (compared to question 1.1) after the 13 questions have been divided into two types, proper care should be taken when placing confidence in the results regarding this main effect; ideally, the study would be repeated with a larger *LTs* sample size.

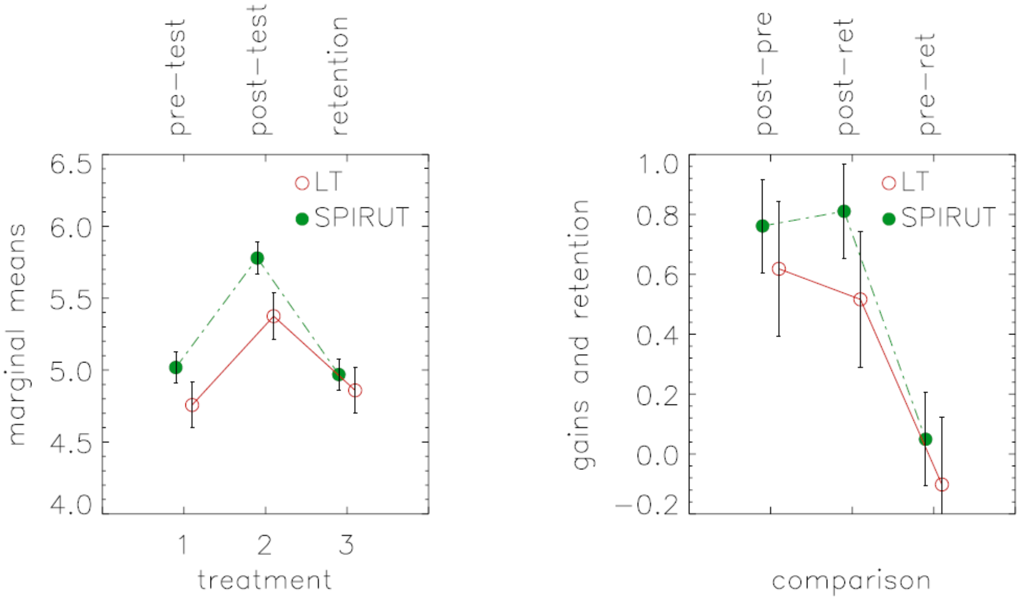


Figure 4.4: Students who worked in the *SPIRUT* group were able to perform simple mathematical calculations and recall basics components of the concept better than students who worked in the *LTs* group left graph). Note however, that the difference in gains between the two groups is only marginal (right graph), suggesting that the indicated main effect is partially due to the difference in means on the pretest, in partial violation of Levene’s assumptions.

Were the treatments successful?

Both groups achieved significant gain as measured on the post-test immediately following treatment (right hand panel Fig. 4.5). Pairwise comparison revealed that while students in the *SPIRUT* group achieved higher scores on the post-test compared to the pre-test, their scores on the retention tests were also significantly lower as compared to the post-test ($M = 4.97$ and $M = 5.77$ respectively, $p = .000$, $p < .0005$). At the same time, students in the *LTs* group also achieved significantly higher scores on the post-test compared to their scores on the pre-test ($M = 5.38$ and $M = 4.76$ respectively, $p = 0.20$, $p < .05$), but their scores on the retention test were not significantly lower than on the post-test ($M = 4.86$ and $M = 5.38$ respectively, $p = .069$, $p > .05$), (Fig. 4.3). Since there was no significant difference between the two treatments in gains/losses between the retention test and pre-test (right hand panel Fig. 4.4) the findings are that there was no benefit from one treatment to the other for long-term knowledge gain (Table 4.7).

Were there differences in gains between the groups?

Post hoc comparison using Tukey's HSD indicated that students who learned the stellar parallax concept by interacting with a computer-based tutorial scored significantly higher on the post-test on *Type I* questions ($M = 5.78$, $SD = 1.22$) compared to students who worked with paper-based tutorials ($M = 5.38$, $SD = 1.46$). In addition, pairwise comparisons tests with Bonferroni adjustment also indicated that the .404 difference between means on the posttest of the two groups was statistically significant ($p = .038$, $p < .05$). On the other hand, a *t*-test demonstrated, *SPIRUT* ($M = .761$, $SD = 1.192$) and

LTs (*M* = .618, *SD* = 1.510), (*t* (197) = 2.723, *p* = .471), that this difference in means did not translate into a significant difference in gains (post-test compared to pre-test), (Table 4.8), between the two groups (right hand panel Fig. 4.4). The significance level of the difference in gains is *p* = 0.471, something which is clearly seen in Figure 4.5. Therefore, it cannot be concluded based upon the present study that there was a significant difference in gains between the treatments.

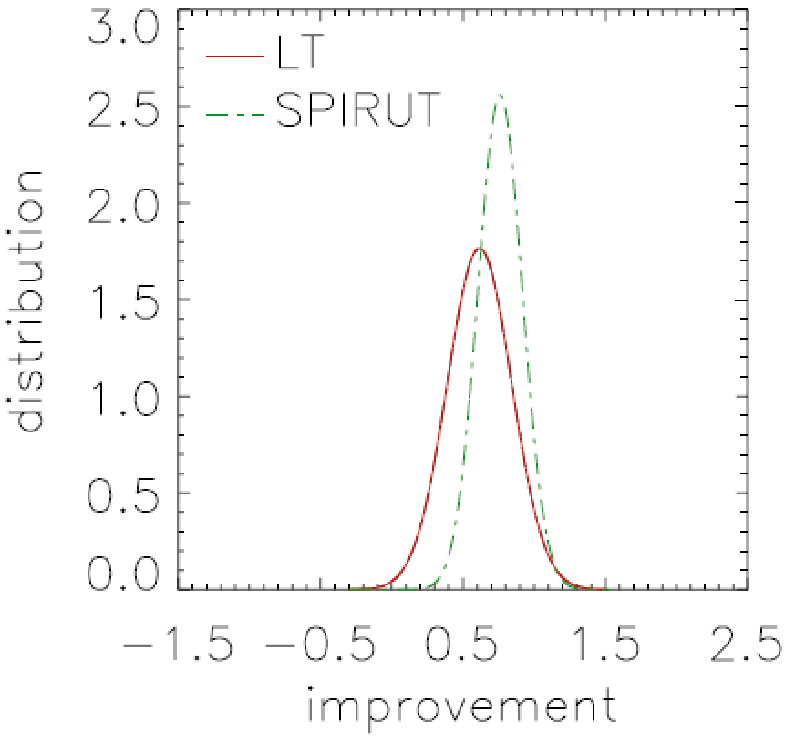


Figure 4.5: While the *SPIRUT* group outperformed the *LT* group in terms of gains, the difference between these two results is not statistically significant.

Table 4.7: *Mean Scores on Type I questions of Stellar Parallax Assessment and Standard Deviations for the SPIRUT and LTs groups on three assessment tests (n = 135 in SPIRUT and n = 64 in LTs)*

	Measurement Time (three assessment tests)			
Student Groups	Pre-test	Post-test	Retention test	Overall
SPIRUT				
M	5.02	5.77	4.97	5.26*
SD	1.08	1.22	1.39	1.29
LTs				
M	4.76	5.38	4.86	5.00*
SD	1.27	1.46	1.37	1.39
Overall				
M	4.93	5.65	4.93	
SD	1.15	1.31	1.38	

Note. * $p < .05$ for tests contrast.

Means in the same row sharing the same letter superscript differ at $p < .05$.

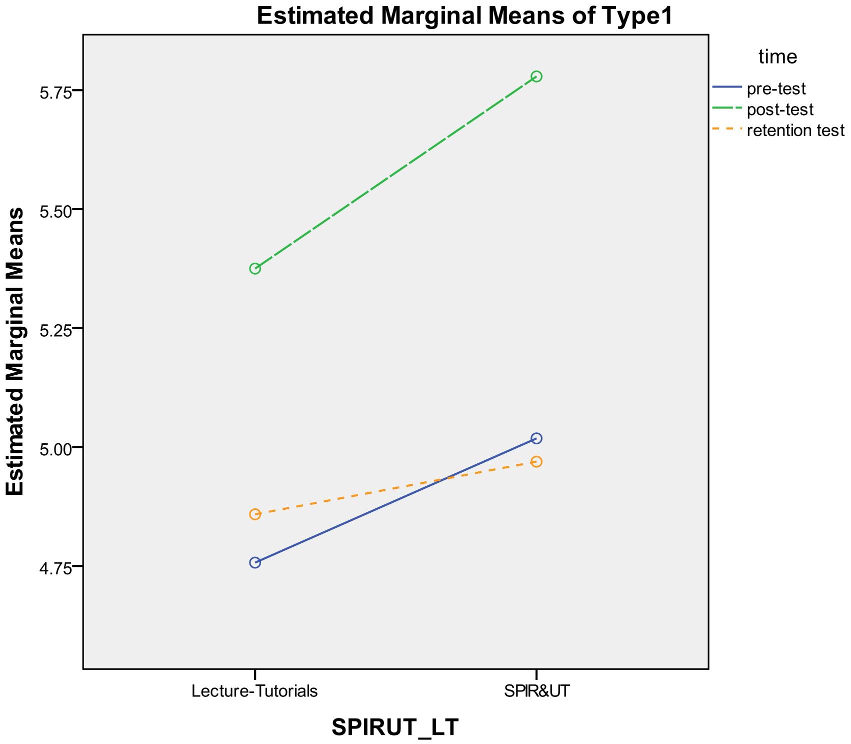


Figure 4.6: Students who worked in *LTs* group retained skills of simple computations better as compared to *SPIRUT* group on the retention test but this improvement is not statistically significant compared to the results on the pre-test of the *LTs* group.

Table 4.8: *The paired-samples t-test on students' gains in SPIRUT and LTs groups*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
SPIRUT	135	.761*	1.192	.0558	.963
Lecture-Tutorials	64	.618*	1.510	.0246	.990

Note. **p* < .05, *p* = 0.471

Summary

Thus, the study revealed that students from both groups improved their mathematical skills (*Type* I) on the test following immediately after their interactions with tutorials. Students who practiced calculations and learned about the concept using computer-based tutorials, *SPIRUT*, achieved significantly higher scores on the test compared to students who practiced their skills using a paper-based tutorial, *LTs*. However, these higher scores are not based on a significant difference in gains (post-test versus pre-test) as in this case the difference in pre-test scores of the two groups accounted for a sizeable fraction of the difference in performance. In addition, students who worked with the paper-based tutorials retained their skills slightly better and after eight weeks later had higher scores (even though not statistically significant) on the retention test as compared to the scores on their pre-test while the computer-based group had returned to their original level without any significant change.

1.3 Type II Knowledge Gain

Do students learn how to answer questions that require them to reason with knowledge better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

Students using computer-based tutorials outperformed their paper-bases peers on pre-test to post-test gains, however the difference did not rise to the 0.05-significance level. Neither treatment resulted in positive measurable effects after eight weeks (Table 4.9).

A 2 (groups) x 3 (assessment tests) factorial ANOVA was performed on students learning achievements on a subset of questions that required *Type II* knowledge. The students learning achievements were measured by this subset of question on the pre-test, post-test, and retention test of the *SPA* instrument.

Table 4.9: *At a glance: significance levels for research question 1.3 pertinent to type of treatment*

Type of Treatment				
Main effect: assessments	**Main effect: treatment**	**Gain *SPIRUT* > 0**	**Gain *LT* > 0**	**Gain *SPIRUT* < Gain *LTs***
$p < .001$	$p = .832$	$P < 0.001$	$p = 1.0$	$p = .007$

Was there a main effect?

Results showed that for *Type II* questions there was no significant main effect of any type of tutorials on students learning achievements across the three assessments ($F(1, 591) = .045$, $p = .832$, $p > .05$) which had no effect size ($\eta_p^2 = .000$). However, there was a significant main effect of students performance on three assessments tests across time

(F(2, 591) = 23.294, p = .000, p < .0005) which was a moderate effect (η_p^2 = .073) (Fig. 4.7).

Were the treatments successful?

Post hoc comparisons using Tukey's HSD indicated that students in the *SPIRUT* group improved their knowledge of the stellar parallax concept on the post-test (M = 2.25, SD = 1.06), which was a statistically significant improvement compared to their mean on the pre-test (M = 1.72, SD = 1.01). This observation was also borne by a pairwise comparisons test using Bonferroni adjustment (p = .000, p < .0005). However, *SPIRUT* students did not retain the constructed *Type II* knowledge and received .803 points lower on the retention test which was a statistically significant difference of means (M = 1.44, SD = .95, p = .000, p < .0005). The *LTs* group did not show a significant improvement on the post-test compared to the pre-test (M = 1.95, SD = 1.00 and M = 2.05, SD = 1.07 respectively, p = 1.000, p > .05). Participants of the *LTs* group also did not retain previously constructed knowledge of the concept and scored on average .690 points lower on the retention test (M = 1.36, SD = 1.07, p = .000, p < .0005), (Table 4.10).

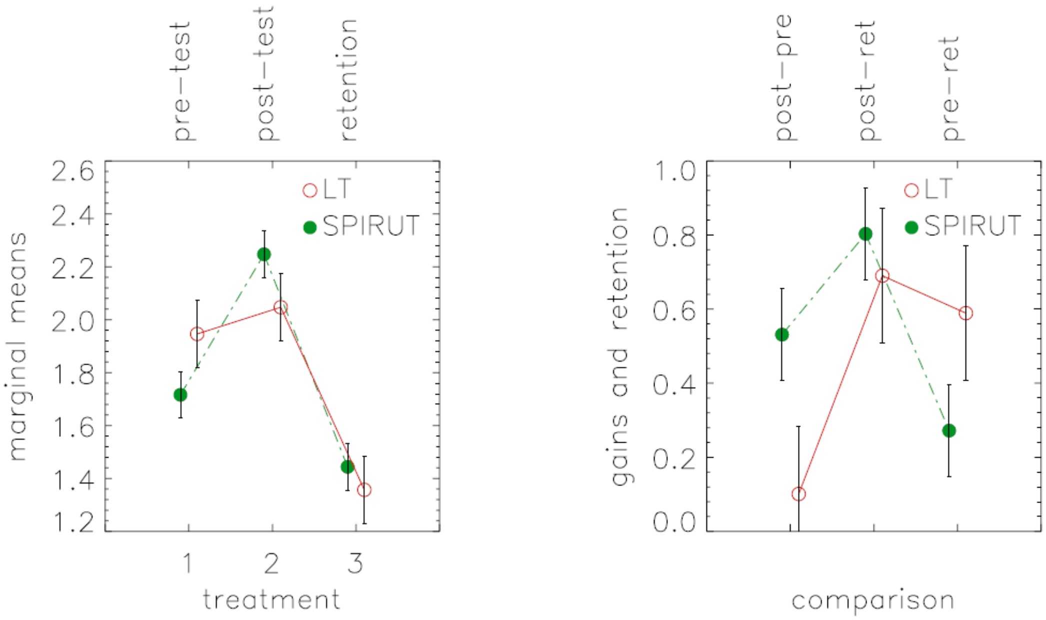


Figure 4.7: The means of the scores of the two groups on the three tests (left panel), and the differences between those means (right panel).

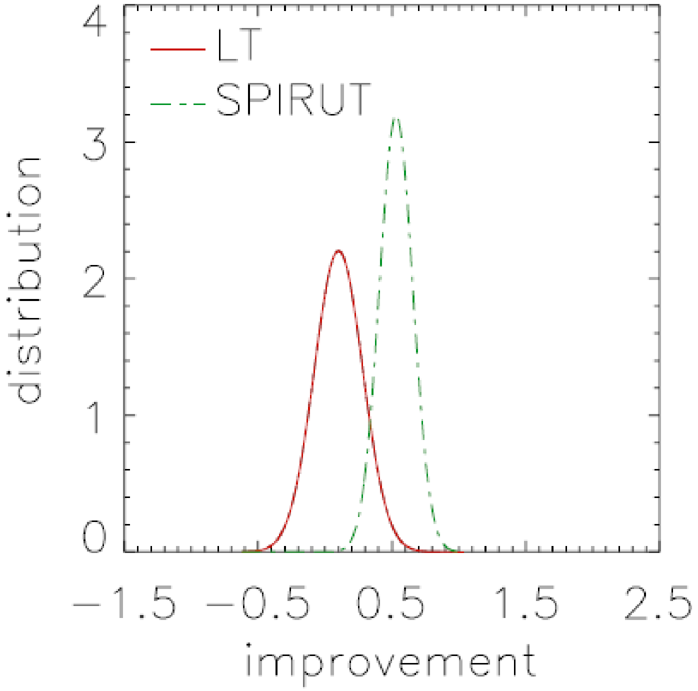


Figure 4.8: While the *SPIRUT* group outgained the *LTs*-group, and while the overlap between the two curves is such that these gains are significant (p = 0.007), it is possible that a fraction of the differences in gain is caused by the *LTs*-group outperforming the *SPIRUT* group on the pre-test (see left panel of Fig. 4.7).

Table 4.10: *Mean Scores on Type II questions of Stellar Parallax Assessment and Standard Deviations for the SPIRUT and LTs groups on three assessment tests (n = 135 and n = 64 in SPIRUT and LTs groups respectively)*

Student Groups		Pre-test	Post-test	Retention test	Overall
SPIRUT					
	M	1.72	2.25	1.44	1.80
	SD	1.01	1.06	.95	1.06
Lecture-Tutorials					
	M	1.95	2.05	1.36	1.78
	SD	1.00	1.11	1.07	1.10
Overall					
	M	1.79	2.18	1.42	
	SD	1.00	1.08	1.00	

Measurement Time (three assessment tests)

Note. * $p < .05$ for tests contrast.

Means in the same row sharing the same letter superscript differ at $p < .05$.

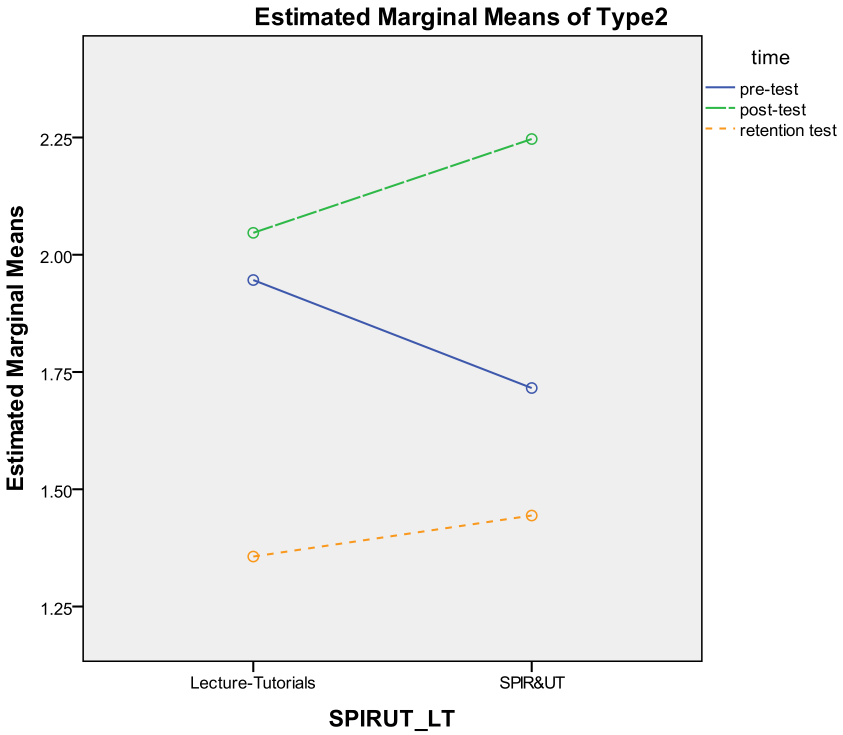


Figure 4.9: Students who worked with computer-based tutorials significantly improved their knowledge of the concept after interactions with *SPIRUT*.

Therefore it can be concluded that overall both treatments had a positive immediate effect on students learning of the concept, but it should be noted that the gain of the lecture tutorial group was not significant, whereas the gain of the *SPIRUT* group was highly significant. Students who worked with *SPIRUT* environments showed a higher level of understanding of the concept by being able to reason with learned basics of the stellar parallax and apply it to answering harder (*Type II*) questions. However, this was not a long-lasting gain since after a six-week period students were not able to achieve the same scores and, some in fact, even scored lower than they did on their pre-test.

Students who worked with the paper-based tutorials were not able to take the learned basics to a higher level and to transfer their knowledge to a hypothetical situation as evaluated on the post-test. Eight weeks later students who worked in the *LTs* group showed a significantly reduced ability in applying the studied concepts to other situations as demonstrated by their scores on the retention test.

Were there differences in gains between the groups?

Even though the results of the paired-samples *t*-test analysis on the posttest versus pre-test gains demonstrated a significant difference ($t(197) = 2.729$, $p = .007$, $p > .05$) in gains in favor of the *SPIRUT* group ($M = .532$, $SD = .994$) (right hand panel Fig. 4.8 and Fig. 4.9), the respective scores on the pre-tests demonstrated that the small group size of the *LTs* group ($M = .101$, $SD = 1.135$) might have resulted in a potential violation of

Leven's assumption. Therefore, the conclusion that the *SPIRUT*-group outperformed the *LTs* group in terms of gains should be held cautiously (Table 4.11).

Table 4.11: *The paired-samples t-test on students' gains in SPIRUT and LTs groups*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
SPIRUT	135	.532*	.994	.364	.701
Lecture-Tutorials	64	.101*	1.135	.179	.381

Note. **p* < .05, *p* = .007

Summary

Similar to the findings for RQ 1.1, students using *SPIRUT* outperformed the comparison group who used the paper based tutorial for *Type II* questions. As before, gains in knowledge were not long-lived. In fact, on these *Type II* questions students performed below their performance on the pre-test measured immediately following the 20 minute classroom lecture.

Overall Summary for Research Questions 1.1, 1.2 and 1.3

Thus, while both kinds of tutorials were effective in helping students improve their knowledge when using a post-test, and the *SPIRUT* was generally superior to the *LTs* group; neither of the treatments was able to help students build a long-lasting understanding of the concept.

II.I.II. Research Question 2

2.1 Learner-controlled and program-controlled SPIRUT

Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions in learner-controlled or program-controlled *SPIRUT*?

The purpose of this analysis was to examine the effects of restrictions on students' interactions in the computer-based tutorial, *SPIRUT*, on constructing knowledge of the stellar parallax concept. In order to investigate this effect on learning, the study participants were randomly assigned to work with one of two versions of *SPIRUT*: learner-controlled, unrestricted environment; or program-controlled version in which students were led by the tutorial in direct response to students' responses to the questions posed in the tutorial (Table 4.12).

It was found that there was no significant difference between two types of interactive tutorials. Students performed significantly better on the post-test as a result of working in the *SPIRUT* environments, and performed equally well in both versions (student or system controlled) of *SPIRUT*.

Table 4.12: *At a glance: significance levels for research question 2.1 pertinent to type of treatment*

Type of Treatment				
Main effect: assessments	**Main effect: learner/program control**	**Gain $UT > 0$**	**Gain $RT > 0$**	**Gain $UT <$ Gain RT**
$p < .001$	$p = .565$	$p < .001$	$p <. 001$	$p = .800$

Was there a main effect?

A multivariate test performed by 2 (groups) x 3 (assessment tests) repeated measures ANOVA determined a significant main effect of the assessment tests ($F_{Wilks' Lambda}$ (2, 132) = 57.373, p < .001), which was a large effect (η_p^2 = .465), (Table 4.13, *Multivariate Test*).

Table 4.13: *Multivariate test: Three-way Repeated Measures ANOVA for Learner-controlled (UT) and Program–controlled (RT) SPIRUT (n = 68 and 67 per cell respectively).*

Source	Wilks' Lambda	*F*	*Hypothesis df*	*Error df*	*P*	Partial eta^2	Power
test	.535	57.373*	2.000	132.000	.000*	.465	1.000
test by *UT_RT*	.999	.90	2.000	132.000	.914	.024	.001

Note. *p < .0005.

Greenhouse-Geisser within subjects test has also revealed that students' scores in the learner-controlled and program-controlled groups were statistically significantly different on three assessments tests (F(2, 260) = 56.676, p = .000, p < 0.001), which was a medium effect (η_p^2 = .299), (Table 4.14, *Within Subjects*).

At the same time, there was no significant difference between students scores in learner-controlled and program-controlled groups (F(1, 133) = .334, p = .565, p > 0.05), which was a medium effect (η_p^2 = .003), (Table 4.14, *Within Subjects*). Students who worked in either one of the two environments performed comparably to each group.

Table 4.14: *Within and Between Groups: Three-way Repeated-Measures ANOVA for* Learner-controll*ed (UT) and Program–controlled (RT) SPIRUT (n = 68 and 67 per cell respectively).*

Source	Type III *SS*	*df*	*MS*	*F*	*p*	Partial eta^2	Power
Between Subjects							
UT_RT	2.267	1	2.267	.334	.565	.003	1.000
Error$_{BS}$	903.965	133	6.797				
Within Subjects							
test	196.906	1.956	100.689	56.676*	.000	.299	1.000
test by *UT_RT*	.337	1.956	.172	.097	.904	.001	.065
Error$_{WS}$	462.077	260.094	1.777				

Note. *$p < .0005$.

However, neither of the main effects depends on the other, as indicated by the non-significant multivariate test for the interaction between groups and tests across time, $F_{Wilks' \ Lambda}$ (2, 132) = .090, $p > .05$, $p = .914$, which was a small effect ($\eta_p^2 = .001$), (Table 4.13, *Multivariate Test*).

Were the treatments successful?

Further scrutiny of the post hoc tests using the Bonferroni correction revealed that students in both groups received statistically significantly higher scores on the post-test ($p = .000$, $p < .0005$) as compared to their scores on pre-test. The difference between means of the pre- and post- tests for *UT* group was 1.255 points and 1.330 points for *RT* group (Table 4.15).

However, scores of the both groups on the retention test were lower and statistically significantly different to their scores on the post-test ($p = .000$, $p < .0005$)

where *UT* lowered its scores by 1.543 points and *RT*'s scores were 1.684 lower compared to the post-test, (Fig. 4.10).

Table 4.15: *Pairwise comparison of two groups: Learner-controlled (UT) and Program–controlled (RT) SPIRUT (n = 68 and 67 per cell respectively)*

groups	(I) test	(J) test	Mean Difference (I-J)	Std. Error	*p*	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
Learner-controlled (*UT*)	1	2	-1.255*	.209	.000	-1.761	-.749
		3	.288	.232	.650	-.274	.850
	2	1	1.255*	.209	.000	.749	1.761
		3	1.543*	.237	.000	.969	2.116
	3	1	-.288	.232	.650	-.850	.274
		2	-1.543*	.237	.000	-2.116	-.969
Program-controlled (*RT*)	1	2	-1.330*	.210	.000	-1.839	-.821
		3	.354	.234	.397	-.213	.920
	2	1	1.330*	.210	.000	.821	1.839
		3	1.684*	.238	.000	.1.106	2.262
	3	1	-.354	.234	.397	-.920	.213
		2	-1.684*	.238	.000	-2.262	-1.106

Note. **p* < .001.

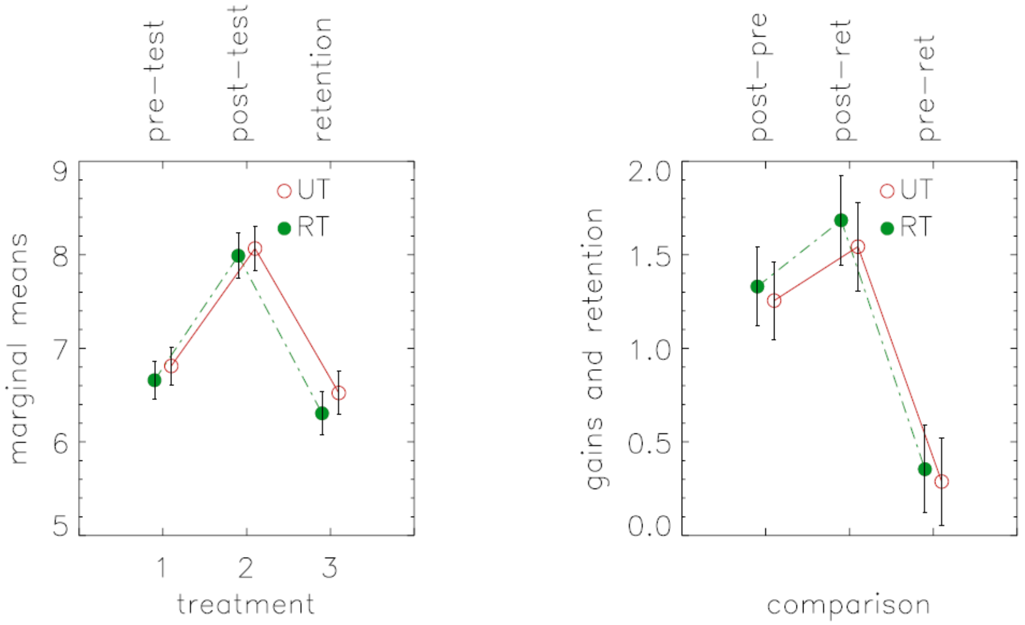


Figure 4.10: While working in two different environments of learner-controlled and program-controlled *SPIRUT*, students in both groups achieved comparable results.

Were there differences in gains between the groups?

Distribution of scores on pre- and post- tests as indicated by the pairwise comparison test revealed - with 95% confidence - that actual gain of *UT* students on the post-test is within .749 and 1.761 points and within the range of .821 and 1.839 points for *RT* students, respectively (Table 4.15). The two distributions almost completely overlap each other. This effect is illustrated in Figure 4.11 below to demonstrate that there is no significant difference between two curves (Fig. 4.11). Thus, students in both groups constructed comparable knowledge.

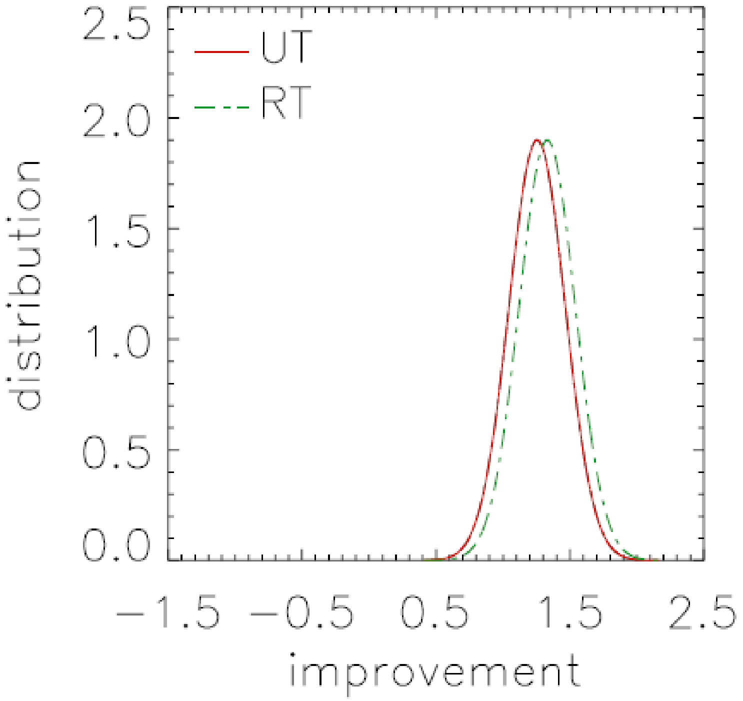


Figure 4.11: The distributions for the gains following treatment for the *UT* group (left curve) and the *RT* group (right curve) approximated by Gaussians based on the standard errors given in Table 4.15. The overlap between these approximate curves put the significance level of the difference at .800.

To verify these findings, the paired-samples *t*-test was conducted to compare students' gains from *UT* and *RT* groups on the post-test. There was no significant difference in the scores for learner-controlled (*M* = 1.255, *SD* = 1.751) and program-controlled (*M* = 1.330, *SD* = 1.688) groups; *t*(133) = .246, *p* = .800, (Table 4.16). These results confirm that there is no difference whether students construct their knowledge while led by a tutorial or have a full control of the environment.

Table 4.16: *The Paired-samples t-Test on Students Gains in Learner-controlled and Program-controlled environments of SPIRUT*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
Learner-controlled	68	1.255	1.751	.838	1.676
Program-controlled	67	1.330	1.688	.922	1.737

Note. $*p < .001$, $p = .800$

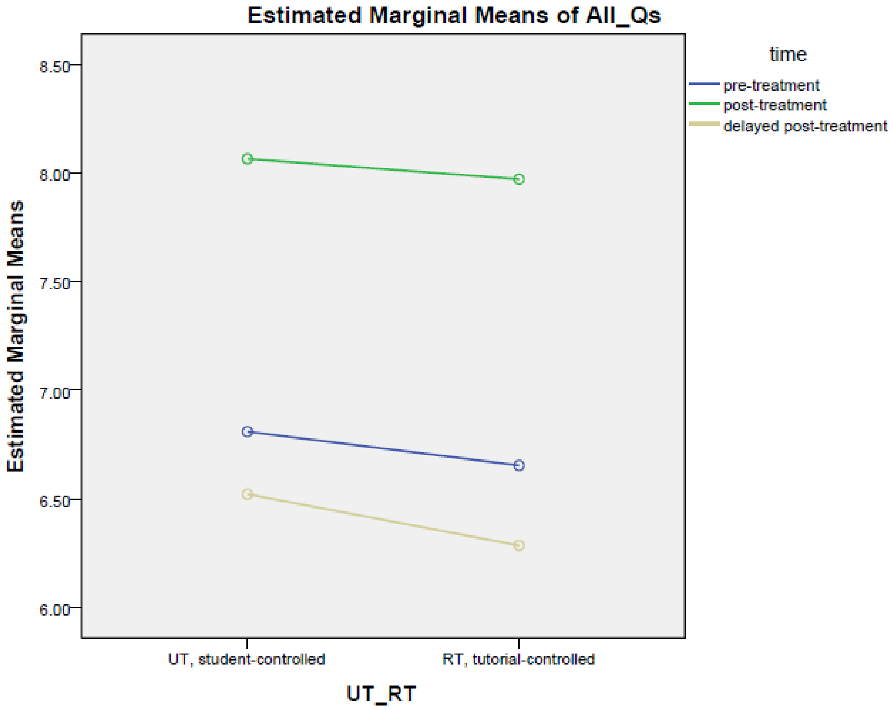


Figure 4.12: This plot was generated by factorial ANOVA to demonstrate means of two groups on three tests.

Summary

We can, therefore, conclude that both versions of the tutorial elicit a statistically significant gain in students' knowledge of the concept as tested immediately after the treatment, and that neither produces a significant effect on retention of the constructed knowledge.

2.2 Type I Knowledge Gains

Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions in learner-controlled or program-controlled *SPIRUT*?

It was found that both treatments were equally beneficial (p = .592) to students' improvement of mathematical skills and memorization of basic components of the stellar parallax concept but only as a temporary effect (Table 4.17).

Another 2 (groups) x 3 (assessment tests) factorial ANOVA was performed to compare two groups of students who worked in *SPIRUT* interactive environment where one group, *Unrestricted Tutorial* (*UT*), was a learner-controlled environment and another group, *Restricted Tutorial* (*RT*), was a program-controlled environment. A follow-up the paired-samples *t*-test was performed to compare students' gains from both groups on the post-test.

Table 4.17: *At a glance: significance levels for research question 2.2 pertinent to type of treatment*

Type of Treatment				
Main effect: assessments	**Main effect: Learner/Program-controlled**	**Gain *UT* > 0**	**Gain *RT* > 0**	**Gain UT < Gain *RT***
$p < .001$	$p = .464$	$p < .001$	$p < .001$	$p = .592$

Was there a main effect?

Results showed that that there was no significant main effect of a type of the tutorial environment on students' knowledge construction (*F(*1, 399) = .536, p = .464, p > .05), which was a small effect (η_p^2 = .001). There was, however, a statistically significant main effect of the assessment tests. Students gained and lost their constructed knowledge as it was shown by their results on the three assessment tests (*F(*2, 399) = 17.105, p = .000, p < .001), which was a large effect (η_p^2 = .083).

Table 4.18: *Tests of Between-Subjects Effects on UT and RT groups*

Source	*SS*	*df*	*MS*	*F*	*p*	Partial eta^2	Observed power
UT_RT	.825	1	.825	.536	.464	.001	.113
Test	67.833	2	27.867	18.105*	.000	.083	1.000
UT_RT * test	.812	2	.406	.264	.768	.001	.092
Error	614.127	399	1.539				

Note. **p* < .001.

Were the treatments successful?

Follow-up with post hoc comparisons using Tukey's HSD and Bonferroni adjustment tests indicated that both environments were equally helpful to students to learn how to perform simple mathematical operation – to find a distance to a star -- and learn basic terms and constructs of the stellar parallax concepts. Mean scores of both groups, *UT* and *RT*, on the post-test (*M* = 5.77, *SD* = 1.25 and *M* = 5.79, *SD* = 1.20 respectively) were statistically significantly different from the mean score on their pre-tests (*M* = 5.06, *SD* = 1.00 and *M* = 4.97, *SD* = 1.16 respectively), (*p* = .000, p < .001), (Table 4.18). Figure 4.13 demonstrates that students' scores were similar across time (Fig. 4.13).

However, neither of the groups retained enough of the developed skills and scored much lower on the retention test eight weeks later (*M* = 4.97, *SD* = 1.39) which was a statistically significant change in means as compared to the post-test (*p* = .000, p < .001) but not significant as compared to the pre-test (*p* = .944, p > .05) (Table 4.19).

Table 4.19: *Mean Scores on Stellar Parallax Assessment and Standard Deviations for the UT and RT Groups on three Tests (n = 68 and n = 67 in Individuals and Peers groups respectively)*

		Measurement Time			
Student Groups		Pre-test	Post-test	Retention test	Overall
Lerner-controlled	M	5.08**	5.77**	5.07**	5.30
(*UT*)	SD	1.00	1.25	1.37	1.22
Program-	M	4.97*	5.79*	4.87*	5.21
controlled (*RT*)	SD	1.16	1.20	1.51	1.36
Overall	M	5.02*	5.78*	4.97*	
	SD	1.08	1.22	1.39	

Note. * *p* < .001, ** *p* < .05

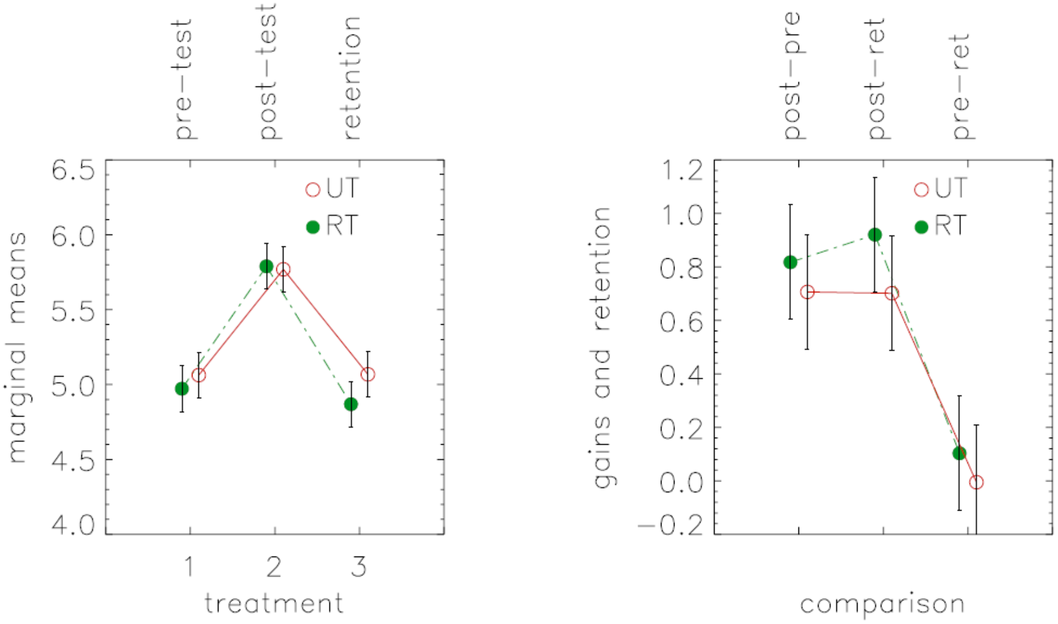


Figure 4.13: Regardless of the type of environments, students in both groups performed identically well on the post-test and comparably lowered their scores eight weeks later.

Were there differences in gains between the groups?

Distribution of scores revealed by the pairwise comparisons test indicated - with 95% confidence - that actual gain on the post-test is within .194 and 1.217 points for the *UT* group and between .301 and 1.332 for the *RT* group, respectively. The two curves representing the distributions of gains almost entirely overlap as shown in figure 4.14, visualizing that one gain is only barely different from the other (Fig. 4.14).

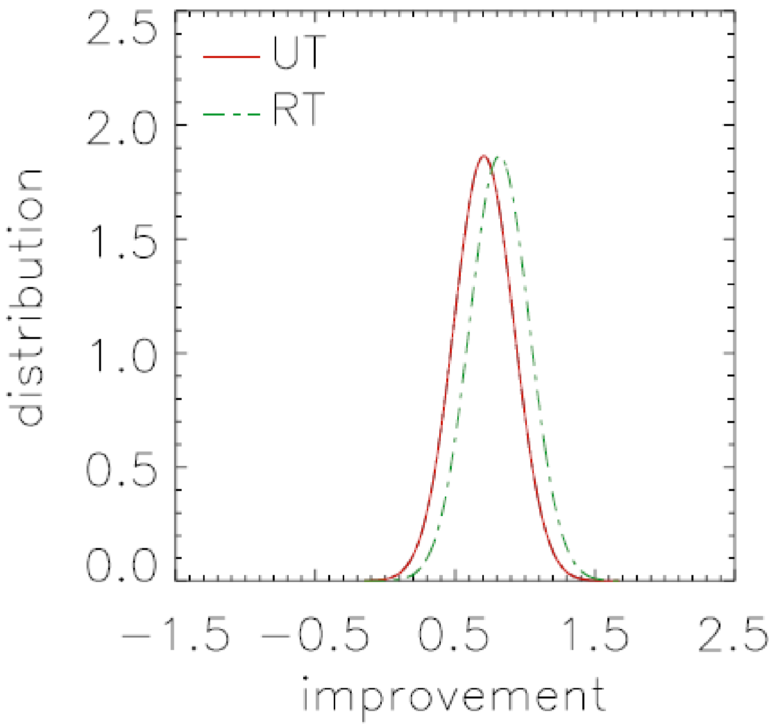


Figure 4.14: The distributions for the gains following treatment for the *UT* group (left curve) and the *RT* group (right curve) approximated by Gaussians based on the standard errors was calculated using groups' means given in Table 4.19. The overlap between these approximate curves put the significance level of the difference at .592.

To verify this finding, the paired-samples *t*-test was performed to compare students' gains from both groups on the post-test. There was no significant difference in the scores for *UT* (*M* = .706, *SD* =1.222) and *RT* (*M* = .816, *SD* = 1.166) groups, (*t*(135) = .538, *p* = .592, *p* > .05). These results suggest that there is no real difference whether students have a full control of the environment or following a tutorial-suggested path, (Table 4.20).

Table 4.20: *The paired-samples t-test on Students Gains in Learner-controlled and Program-controlled environments of SPIRUT*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
Lerner-controlled	68	.706	1.222	.414	.990
Program-controlled	67	.816	1.166	.536	1.097

Note. *p < .001, p = .592

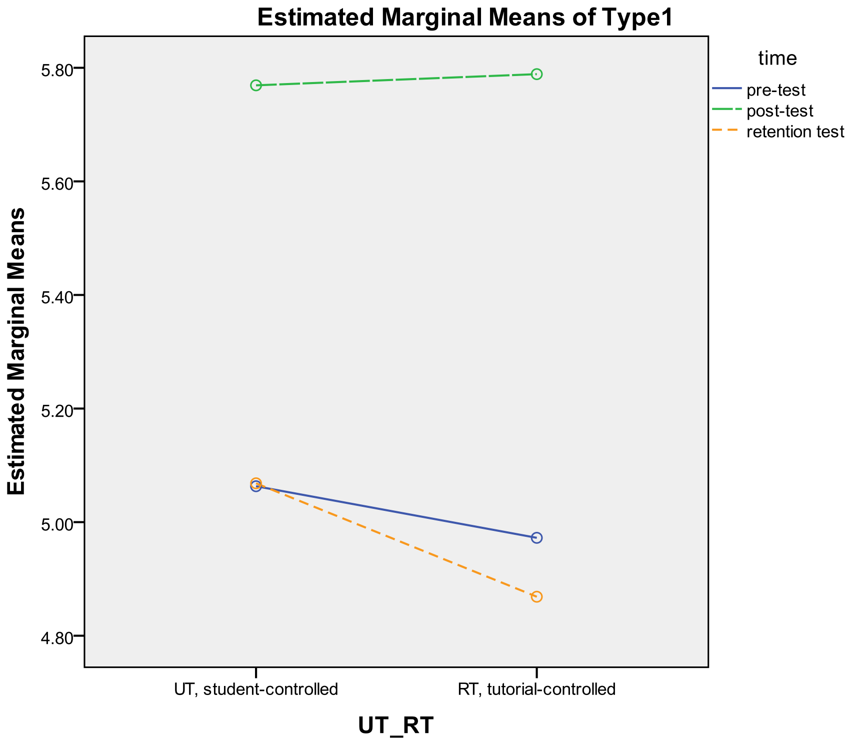


Figure 4.15: This plot was generated by a factorial ANOVA to demonstrate that students in both groups had similar gains and lose in the mean score on the post- and retention tests.

Summary

Thus, it can be concluded that students significantly improved their mathematical skills and memorized important aspects of the concept regardless of whether their interactions with the tutorials were controlled by them or restricted by the tutorial's environment. Although students in both groups performed significantly better on the post-test, the fact that their mathematical skills as measured by their scores on the retention test dropped significantly indicates that the students' achievements were temporary and that neither of the environments helped them to retain their acquired skills and memorized facts.

2.3 Type II Knowledge Gain

Do students learn how to answer questions that require them to reason with knowledge better when their learning processes are aided by interactions in learner-controlled or program-controlled environments of *SPIRUT*?

It was found that independently whether students had a full control of the environment or followed an adaptive path created by the tutorial based on students performances, the participants demonstrated their understanding of the concept and transfer knowledge of the concept answering equally well (p = .982) on research *Type II* subset questions (Table, 4.21).

To answer the research question a 2 (groups: *UT* and *RT*) x 3 (assessment tests) factorial ANOVA was performed on the students' performance on *Type II* questions, a set of harder questions that require use of reasoning skills.

Table 4.21: *At a glance: significance levels for Research Question 2.3 pertinent to type of treatment*

Type of Treatment				
Main effect: assessments	**Main effect: learner/Program-control**	**Gain *UT* > 0**	**Gain *RT* > 0**	**Gain *UT* > Gain *RT*; Gain *UT* < Gain *RT***
$p < .001$	$p = .733$	$p < .001$	$p <. 001$	$p = .826$

Was there a main effect?

Results showed that overall there was no statistically significantly difference between two groups of students ($F(1, 399) = .117$, $p = .733$, $p > .05$), which was unnoticeable effect ($\eta_p^2 = .000$). There was a significant main effect of students' performance on the assessments tests ($F(2, 399) = 43.862$, $p = .000$, $p < .001$) which was a small effect ($\eta_p^2 = .180$), (Table 4.22).

Table 4.22: *Tests of Between-Subjects Effects on UT and RT groups*

Source	*SS*	*df*	*MS*	*F*	*p*	Partial eta^2	Observed power
UT_RT	.144	1	.144	.117	.733	.000	.063
Test	108.082	2	54.041	43.862*	.000	.180	1.000
UT_RT * test	.335	2	.167	.136	.873	.001	.071
Error	491.593	399	1.232				

Note. *$p < .001$.

Were the treatments successful?

Pairwise comparisons post hoc tests using Tukey's HSD and Bonferonni adjustments revealed that whether students used a learner-controlled or a program-controlled environment to learn about the concept, their abilities of applying *Type II* knowledge was comparable. Both groups of students showed statistically significant improvement on the post-test (*M* = 2.94, *SD* = 1.16, $p = .000$, $p < .001$). However, eight weeks later both groups of students performed poorly in finding correct answers using their reasoning skills as shown by their significantly lower mean scores on the retention test (*M* = 1.69, *SD* =1.03, $p = .000$, $p < .001$) (Table 4.23).

Table 4.23: *Mean Scores on Stellar Parallax Assessment and Standard Deviations for the UT and RT Groups on three Tests (n = 68 and n = 67 in Individuals and Peers groups respectively)*

Measurement Time Student Groups		Pre-test	Post-test	Retention test	Overall
Learner-controlled	M	2.17*	2.94*	1.75	2.28
(*UT*)	SD	1.22	1.10	.90	1.19
Program-controlled	M	2.18*	2.94*	1.63*	2.25
(*RT*)	SD	1.02	1.22	1.15	1.25
Overall	M	2.18*	2.94*	1.69	
	SD	1.13	1.16	1.03	

Note. * $p < .001$, ** $p < .05$

In addition, mean scores of students who worked with program-controlled *SPIRUT* were statistically significantly lower on the retention test (*M* = 1.63, *SD* = 1.15) than their scores on the pre-test (*M* = 2.18, *SD* =1.02, $p = .013$, $p < .05$). At the same time, mean scores of students who worked in a restriction-free learner-controlled environment of *SPIRUT* were lower on the retention test but not statistically significantly different than their mean scores on the pre-test (*M* = 1.75, *SD* = .90 and *M* = 2.17, *SD* = 1.22 respectively, *p* = .078, p > .05) (Fig. 4.16).

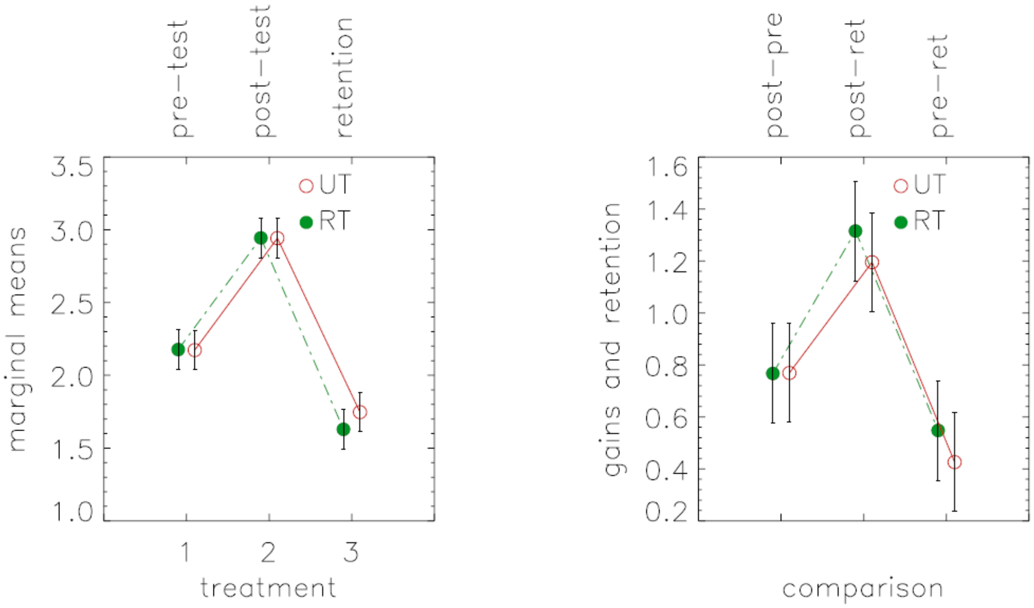


Figure 4.16: Regardless of the type of environments, students in both groups achieved similar results on the post-test and comparably lowered their scores eight weeks later.

Were there differences in gains between the groups?

Distribution of scores revealed by the pairwise comparisons test indicated - with 95% confidence - that actual gain on the post-test is within .312 and 1.227 points for the *UT* group and between .305 and 1.227 for the *RT* group, respectively. The two curves for the distributions of means in figure 4.17 visibly overlap (Fig. 4.17). In fact, it is not possible to find more identical outcomes within the restrictions of the study.

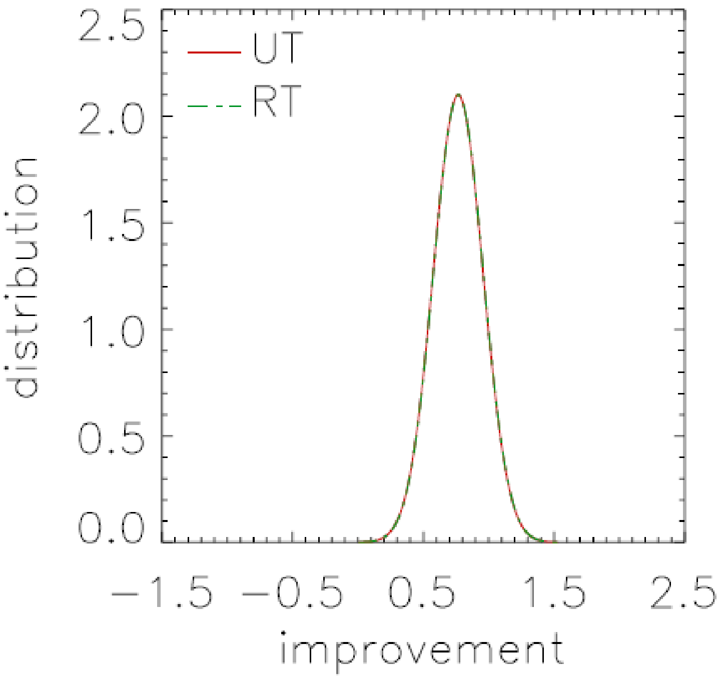


Figure 4.17: This plot of the distributions of means between *UT* and *RT* groups is a visual demonstration that there is virtually no difference in means between these groups.

To verify and further the results of the ANOVA model, the paired-samples t-test was conducted to compare gains in score means of the two groups. There was no significant difference in the scores for *UT* (*M* = .551, *SD* = 1.102) and *RT* (*M* = .513, *SD* = .879) groups, (*t*(133) = .220, *p* = .826), (Table 4.24). These results suggest that whether students work at their own pace and select their own way of learning the concept or whether they follow a tutorial-determined learning procedure; they develop comparable reasoning skills and ability of transferring their knowledge to unknown situations.

Table 4.24: *The paired-samples t-test on Students Gains in Learner-controlled and Program-controlled environments of SPIRUT*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
Lerner-controlled	68	.551	1.102	.287	.815
Program-controlled	67	.513	.879	.302	.725

Note. **p* < .05, *p* = .982

Summary

It can be concluded that, regardless of whether students had full control over the environment and were determining how they constructed their knowledge, or whether the environment was taking its users through a step-by-step process, students acquired comparable levels of *Type II* knowledge. Although both groups showed significant improvement in their scores on the post-test, six weeks later their ability to apply their knowledge was significantly lower compared to what it was on the post-test. In fact, both groups dropped to below their pre-test levels with the largest differences observed in the program-controlled group.

Overall Summary for Research Questions 2.1, 2.2 and 2.3

There is no difference in any of the learning outcomes between restricting the tutorial and between leaving it up to students to navigate their way through the *SPIRUT* system.

II.I.III. Research Question 3

3.1 Peers and Individuals

Do students construct greater knowledge of the stellar parallax concept when their learning processes are supported by working with a partner or working independently?

Surprisingly, the results showed the opposite of the researcher's expectation. Namely, it was found that individuals outperformed their paired peers. This finding while not significant (p = .146) was a substantial difference in a counter-intuitive direction, and therefore, the qualitative part of the research, to be described later on in this Chapter, was used to examine whether this surprising finding could be explained (Table 4.25).

A repeated measures 2 (peerss and individuals) x 3 (assessment tests) ANOVA was performed to investigate whether there were significant differences in students performances on three assessment tests depending on whether they worked together or individually. Such differences would indicate that the interaction during treatments had an effect on students' conceptual change and whether this effect, if present, is manifest in the students' understanding of the stellar parallax concept.

An independent sample *t*-test was performed to investigate whether gains in points within and between the differently paired groups of students were significant different. If a difference was found, that would indicate that the interaction during treatments indeed created a gain in conceptual understanding and helped students to further their knowledge of the stellar parallax concept.

Table 4.25: *At a glance: significance levels for research question 3.1 pertinent to peer instruction*

Type of Treatment				
Main effect: assessments	**Main effect: *Peers/individuals***	**Gain *Peers* > 0**	**Gain *Individuals* > 0**	**Gain *Peers* < Gain *Individuals***
p = 000	p = .048	$p < .001$	$p = .001$	$p = .146$

Was there a main effect?

The multivariate test indicated a statistically significant different main effect of the three assessment test ($F_{Wilks' Lambda}$ (2, 196) = 67.341, p = .000, $p < .001$), which was a strong effect (η_p^2 = .407) that accounted for 41 percent of variance. Students performed significantly different across time as it was indicated by the statistically significantly different means on three assessment tests (Table 4.26, *Multivariate Test*).

Table 4.26: *Multivariate test: Three-way Repeated-Measures ANOVA for Individuals and Peers on three assessment test, (n = 99 and 100 per cell respectively).*

Source	Wilks' Lambda	*F*	*Hypothesis df*	*Error Df*	*p*	Partial eta^2	Power
test	.593	67.341*	2.000	196.000	.000*	.407	1.000
test by Indiv._Peers	.972	2.836	2.000	196.000	.061	.028	.552

Note. *$p < .0005$.

In accordance with Willks' test, Greenhouse-Geisser statistics test results indicated that mean of three assessment tests reflecting students performance differed statistically significantly across time (F(2, 394) = 66.786, p = .000, $p < .0005$), which was a moderate effect (η_p^2 = .253), (Table 4.27, *Within Subjects*).

Table 4.27: *Within and Between Groups: Three-way Repeated-Measures ANOVA for Individuals and Peers on three assessment test, (n = 99 and 100 per cell respectively).*

Source	Type III *SS*	*df*	*MS*	*F*	*p*	Partial eta^2	Power
Between Subjects							
Individuals_Peers	28.174	1	28.174	3.948**	.048	.020	.507
Error$_{BS}$	1406.022	197	7.137				
Within Subjects							
test	236.432	1.977	119.591	66.786*	.000	.253	1.000
test by Indiv._Peers	10.342	1.977	5.231	2.921	.056	.015	.565
Error$_{WS}$	697.411	394	1.770				

Note. *p < .0005, **p < .05.

In addition, Bonferroni adjustment post hoc test was performed on pairwise comparison of two groups on combined tests scores from three assessments. The tests revealed that individuals group statistically significantly outperformed pairs group as measured by overall means difference of .434 points ($p = .048$, $p < .05$), (Table 4.27, *Between Subjects*).

The Greenhouse-Geisser test performed on groups of students and the assessments tests revealed only a marginally significant main effect of groups on the three assessments ($F(2, 394) = 2.921$, $p = .056$, $p > 0.05$) with a small effect size ($\eta_p^2 = .015$), (Table 4.13, *Within Subjects*). Also, Wilks' multivariate test indicated a weak interaction ($F_{Wilks'\ Lambda}(2, 196) = 2.836$, $p = .061$, $p > .05$), which was a moderate effect size ($\eta_p^2 = .028$) and had insufficient power (observed power, .552), (Table 4.26, *Multivariate Test*).

Thus, even though there is an indication of marginal dependency of one variable on another, the modest effect size and insufficient statistical power of both variables preclude us from drawing a strong conclusion but rather suggest further investigation of each of the variables.

Were the treatments successful?

To examine how students' scores changed over time in relation to one another, post hoc comparison using the Bonferroni adjustment test was performed. The test revealed that students who independently gained 1.293 points on the post-test while students who worked in peers gained .925 points, both these differences in means were statistically significant ($p = .000$, $p < .0005$). However, eight weeks later students who worked independently reduced their mean scores by 1.344 points as compared to their post-test results ($p = 000$, $p = .0005$) which was also a mere .052 points as compared to the pre-test which was an insignificant change ($p = 1.000$, $p > .05$). Students who worked in peers lost 1.619 points on the retention test when compared to the post-test ($p = .000$, $p < .0005$) which also was .694 points lower as compared to the pre-test which was a statistically significantly difference ($p = .001$, $p < .0005$), (Table 4.28). The left panel of figure 4.18 demonstrates that students' performances on post-test and retention tests were different where individuals consistently received higher scores than the peer group (Fig. 4.18).

Thus, it can be suggested that students who worked individually indeed performed better on three assessment tests. In order to find out whether these different performances resulted in greater knowledge the data has to be examined further and differences in gains between two groups need to be compared.

Table 4.28: *Pairwise comparison of two groups: Individuals and Peers, (n = 99 and n = 100 per cell respectively) on three assessment tests*

groups	(I) test	(J) test	Mean Difference (I-J)	Std. Error	*p*	95% Confidence Interval for Difference	
						Lower Bound	Upper Bound
Individuals	1	2	-1.293*	.179	.000	-1.724	-.861
		3	.052	.194	1.000	-.417	.520
	2	1	1.293*	.179	.000	.861	1.724
		3	1.344*	.194	.000	.875	1.813
	3	1	-.052*	.194	1.000	-.520	.417
		2	-1.344*	.194	.000	-1.813	-.875
Peers	1	2	-.925*	.178	.000	-1.354	-.496
		3	.694	.193	.001	.228	1.160
	2	1	.925*	.178	.000	.496	1.354
		3	1.619*	.193	.000	1.153	2.086
	3	1	-.694	.193	.001	-1.160	-.228
		2	-1.619*	.193	.000	-2.086	-1.153

Note. *$p < .001$, ** $p < .05$.

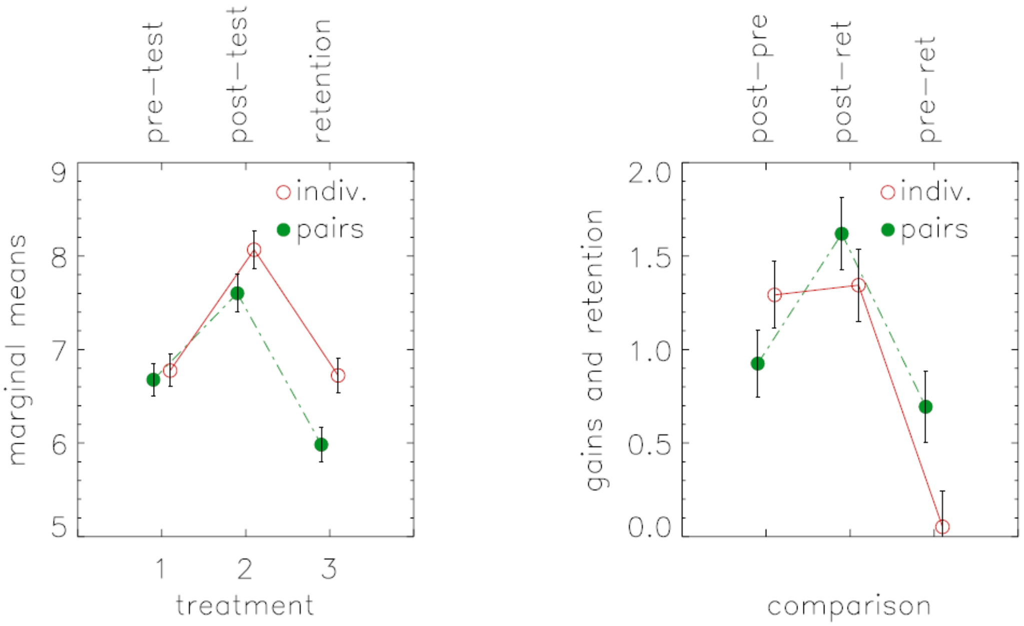


Figure 4.18: Students achievements on three assessment tests (left panel); Students who worked with peers showed significant decrease in their knowledge of the concept (right panel).

<u>Were there differences in gains between the groups?</u>

Distribution of scores as revealed by the pairwise comparison test indicated - with 95% confidence - that actual gain on the post-test is within .861 and 1.724 points for the Individuals group and between .496 and 1.354 for the Peers group, respectively (Table 4.28). The small overlap of two distributions shown in figure 4.19 visualizes that one gain is marginally different from the other (Fig. 4.19).

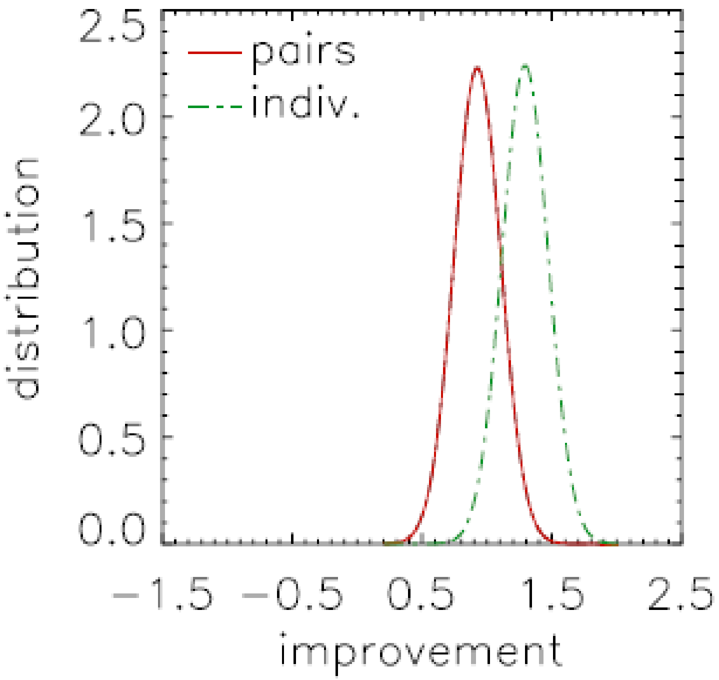


Figure 4.19: The distributions for the gains following treatment for the Individuals group (left curve) and the Pairs group (right curve) approximated by Gaussians based on the standard errors given in Table 4.28. The overlap between these approximate curves put the significance level of the difference at 0.146.

The paired-samples *t*-test was conducted to compare gains of individuals and peer groups. There was no significant difference in gain found for individuals (*M* = 1.293, *SD* = 1.751) and peers (*M* = .925, *SD* = 1.803) groups (*t*(197) = 1.460, *p* = .146, *p* > .05), (Table 4.29). These results suggest while students who worked individually consistently performed better than students who worked with peers, their gains as indicated by means on the post-test and pre-test are still comparable (Fig. 4.20).

Table 4.29: *The paired-samples t-test on students' gains in the Individuals and Peers groups*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
Individuals	99	1.293	1.751	.946	1.641
Peers	100	.925	1.803	.570	1.281

Note. $^*p < .05$, $p = .146$

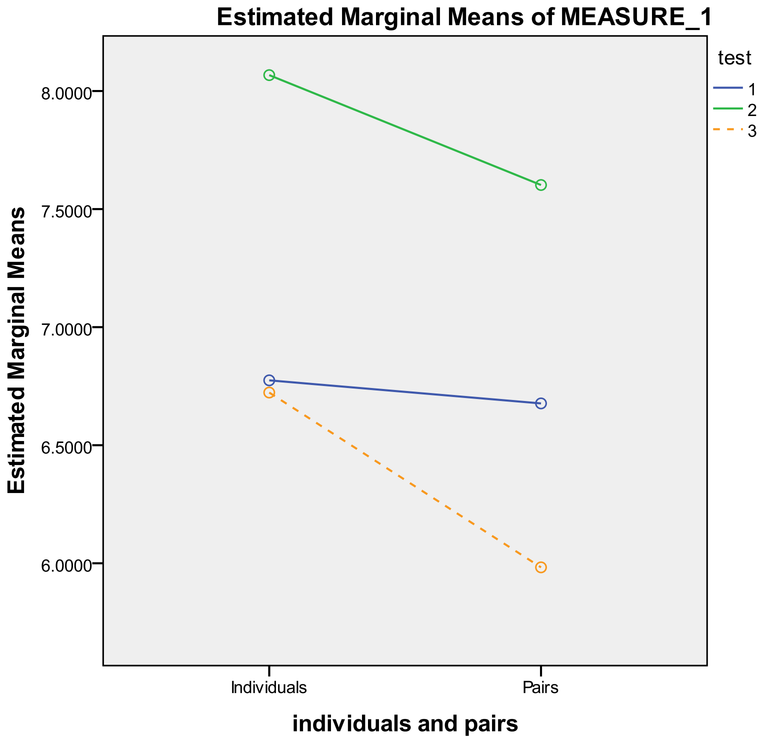


Figure 4.20: This plot was generated by factorial ANOVA to demonstrate means of two groups on three tests. Students who worked independently received higher scores on the post-test and retained the learned material better compared to students who worked in peers.

Summary

Therefore, it can be concluded that while students who worked independently or in pairs constructed comparable knowledge of the concept, students who worked independently retained their knowledge better. In addition, there was a clear tendency revealed that students who work individually consistently performed better on post- and retention assessment tests.

3.2 Type I Knowledge Gains

Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are supported by working with a partner or working independently?

It was found that while both groups of students had similar gain on the post-test (p = .338), individuals consistently received higher scores and significantly outperformed peer groups on the retention test, (Table 30).

To answer the question a 2 (individuals and peers) x 3 (assessment tests) factorial ANOVA was performed on students' performance on three assessment test with selected *Type I* questions. To find out whether students achieved significantly higher gains as a result of their work individually or with peers, their post-test scores were statistically evaluated by the paired-samples *t*-test.

Table 4.30: *At a glance: significance levels for research question 3.2 pertinent to type of interaction*

Type of Interaction				
Main effect: assessments	**Main effect: *Peers*/*Individuals***	**Gain *Peers* > 0**	**Gain *Individuals* > 0**	**Gain *Peers* < Gain *Individuals***
$p < .001$	$p = .001$	$p < .001$	$p = .065$	$p = .388$

Was there a main effect?

There was a significant main effect of whether students worked independently or with peers on their overall performance on the assessment tests ($F(1, 579) = 10.401$, $p = .001$, $p < .05$), which was a small effect ($\eta_p^2 = .017$), (Table 4.31). There was also a significant main effect of students overall performance across time ($F(2, 579) = 20.863$, $p = .000$, $p < .001$), which was a large effect ($\eta_p^2 = .066$), (Fig.4.21).

Table 4.31: *Tests of Between-Subjects Effects on Individuals and Peers groups*

Source	*SS*	*df*	*MS*	*F*	*p*	Partial eta^2	Observed power
Indiv_Peers	16.908	1	16.908	10.401**	.001	.017	.896
Test	67.833	2	33.917	20.863*	.000	.066	1.000
Indiv_Peers * test	2.671	2	1.335	.821	.440	.003	.191
Error	960.758	591	1.626				

Note. $*p < .001$, $** p < .05$.

Were the treatments successful?

Further investigation using multiple comparisons using Tukey’s HSD test revealed that overall both groups statistically significantly improved on the post-test as compared to their mean scores on the pre-test (*M* = 4.93, *SD* = 1.38 and *M* = 5.65, *SD* = 1.31, respectively, *p* = .000, *p* < .001). However, eight weeks later on the retention test, students' mean scores in both groups were statistically lower than their mean scores on the post-test (*M* = 4.93, *SD* = 1.38 and *M* = 5.65, *SD* = 1.31, respectively, *p* = .000, *p* < .001), (Table 4.32).

In addition, post hoc testing using Tukey's HSD and Bonferroni adjustment for multiple comparisons indicated that on the post-test students who worked independently either with computer-based or paper-based tutorials performed marginally better than students who worked with peers (*M* = 5.82, *SD* = 1.35 and *M* = 5.48, *SD* = 1.25, respectively, *p* = .065, *p* < .01). At the same time, on the retention test the mean score of the individuals group on *Type I* questions was statistically significantly higher that the mean score of the peers group (*M* = 5.19, *SD* = 1.37 and *M* = 4.68, *SD* = 1.36 respectively, *p* = .006, p < .05), (Fig. 4.21).

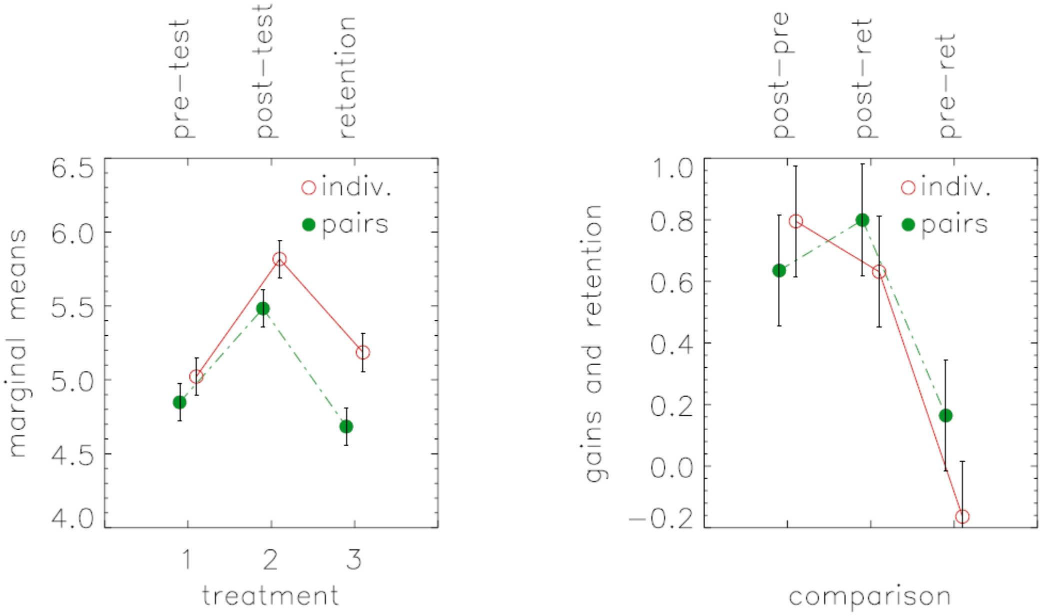


Figure 4.21: Students achievements on three assessment tests (left panel); Students who worked with peers showed significant decrease in their knowledge of the concept (right panel).

Table 4.32: *Mean Scores on Stellar Parallax Assessment and Standard Deviations for the Individuals and Peers groups on three tests (n = 99 and n = 100 in Individuals and Peers groups respectively)*

		Measurement Time			
Student Groups		Pre-test	Post-test	Retention test	Overall
Individuals					
	M	5.02	5.82	5.19**	5.34
	SD	1.12	1.35	1.37	1.33
Peers					
	M	4.85	5.48	4.68**	5.00
	SD	1.17	1.25	1.36	1.31
Overall					
	M	4.93*	5.65*	4.93*	
	SD	1.15	1.31	1.38	

Note. * $p < .001$, ** $p < .05$.

Were there differences in gains between the groups?

Distribution of scores revealed by the pairwise comparisons test indicated - with 95% confidence - that actual gain on the post-test for *Type I* questions is within .163 and 1.427 points for the Individuals group and between .003 and 1.267 for the peers group, respectively. The sizeable overlap of two distributions shown in figure 4.22 visualizes that one gain is only marginally different, at most, from the other (Fig. 4.22).

As a part of the investigation, the paired-samples *t*-test was conducted to compare students gains on the post-test so to answer a question whether students social engagement help them to sharpen mathematical skills and memorize a number of facts pertained to the learned concept. There was no significant difference in the scores for *Individuals* (*M* = .795, *SD* =1.368) and *Peers* (*M* = .636, *SD* = 1.231) groups, (*t*(133) = 2.55, *p* = .388, p > .05), (Table 33). These results suggest that there is no real difference whether students perform their task individually or with peers.

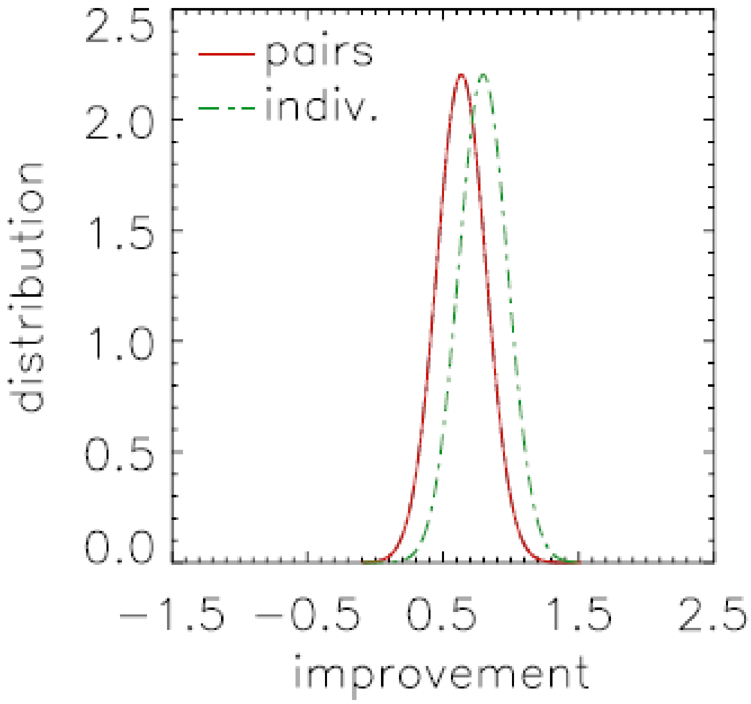


Figure 4.22: The distributions for the gains following treatment for the Individuals group (left curve) and the *Peers* group (right curve) approximated by Gaussians based on the standard errors was calculated using groups' means given in Table 4.32. The overlap between these approximate curves put the significance level of the difference at .388.

At the same time, visual representation of the students' performance on all tests shows that students who worked individually consistently outperform students who worked with partners (Fig. 4.23).

Table 4.33: *The Paired-samples t-Test on Students Gains in Individuals group and in Peers group*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
Individuals	99	.795	1.368	.524	1.066
Peers	100	.636	1.231	.393	.878

Note. *$p < .05$, $p = .388$

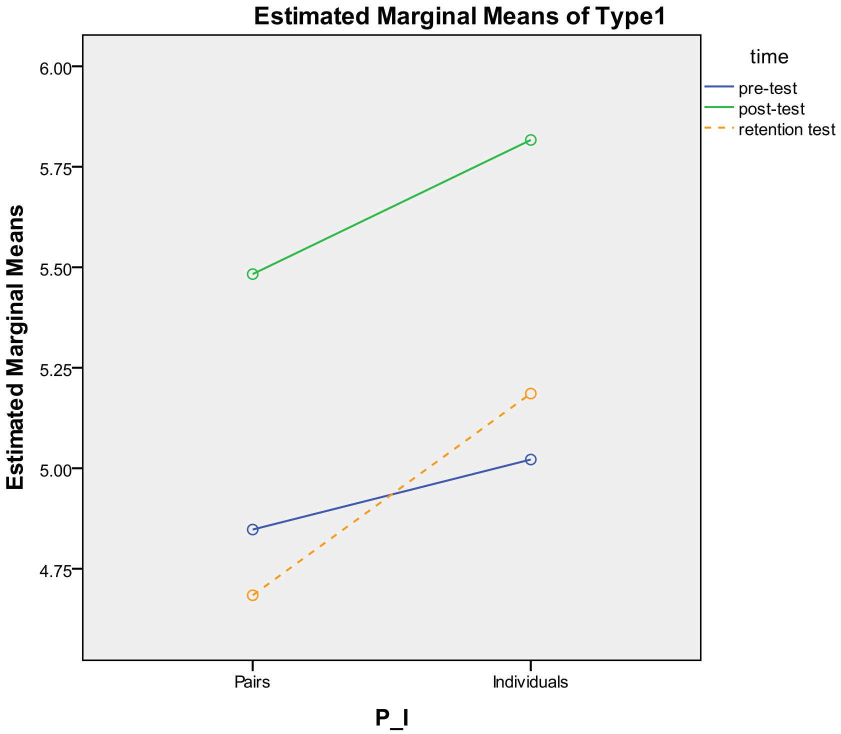


Figure 4.23: This plot was generated by a factorial ANOVA to demonstrate that students in both groups successfully memorized constituents of the stellar parallax concept and were able to perform simple calculations while students who worked independently also retained the acquired skills better.

Summary

Therefore, it can be concluded that independent of whether students, using different tutorials, worked with peers or worked individually, that they were able to sharpen their mathematical skills and learn more about the stellar parallax concept. At the same time, students who worked independently were able to achieve higher scores on the post-test thus indicating that working independently was more beneficial for learning. In addition, eight weeks later students who worked independently remembered more information about the concept than students who worked with peers.

3.3 Type II Knowledge Gains

Do students learn how to reason their answers and transfer knowledge better when their learning processes are supported by working with a partner or working independently?

A 2 (peers and individuals) x 3 (assessment tests) factorial ANOVA was performed on comparison of two groups of students who worked individually or with peers while learning the stellar parallax concept. The paired-samples *t*-test was performed to ascertain the differences in gains between the two groups, (Table 4.34).

Table 4.34: *At a glance: significance levels for Research Question 3.3 pertinent to type of interaction*

Type of Interaction				
Main effect: assessments	**Main effect: *Peers/Individuals***	**Gain *Peers* > 0**	**Gain *Individuals* > 0**	**Gain *Peers* < Gain *Individuals***
$p < .001$	$p = .243$	$p > .05$	$p > .05$	$p = .165$

Was there a main effect?

Results show that overall there was no significant main effect on learning *Type II* knowledge of whether students worked individually or with peers as indicated by their mean scores on the assessment tests ($F(1, 591) = 1.365$, $p = .243$, $p > .05$), which was a small effect ($\eta_p^2 = .002$). However, there was a significant main effect of time on students performance ($F(2, 591) = 27.895$, $p = .000$, $p < .001$), which was a large effect ($\eta_p^2 = .086$), (Table 4.35).

Table 4.35: *Tests of Between-Subjects Effects on Individuals and Peers groups*

Source	*SS*	*df*	*MS*	*F*	*p*	Partial eta^2	Observed power
Indiv_Peers	16.908	1	16.908	10.401**	.001	.017	.896
Test	67.833	2	33.917	20.863*	.000	.066	1.000
Indiv_Peers * test	2.671	2	1.335	.821	.440	.003	.191
Error	960.758	591	1.626				

Note. $^*p < .001$, $^{**}\ p < .05$.

Were the treatments successful?

Post hoc multiple comparisons using Tukey's HSD test revealed that overall differences between mean scores of students in both groups differed significantly across time. Students in both groups significantly improved their scores on the post-test when compared to the pre-test (*M* = 1.79, *SD* = 1.01 and (*M* = 2.18, *SD* = 1.08 respectively, *p* = .000, p < .001), (Table 4.36). Also students scores were significantly lower on the retention test (*M* = 1.42, *SD* = .99) when compared to the post-test (*p* = .000, *p* = .001) and significantly lower when compared to the pre-test (*p* = .001, *p* = .001) indicating that either of the treatments were successful at achieving a long-term effect on the students' *Type II* knowledge gain (Fig. 4.24).

Table 4.36: *Mean Scores on Stellar Parallax Assessment and Standard Deviations for the Individuals and Peers groups on three tests (n = 99 and n = 100 in Individuals and Peers groups respectively)*

		Measurement Time			
Student Groups		Pre-test	Post-test	Retention test	Overall
Individuals					
	M	1.75	2.25	1.54	1.85
	SD	1.05	1.12	1.05	1.11
Peers					
	M	1.83	2.12	1.30	1.75
	SD	.96	1.04	.91	1.03
Overall					
	M	1.79	2.18	1.42	
	SD	1.01	1.08	.99	

Note. * $p < .001$ for tests contrast.

Means in the same row sharing the same letter superscript differ at $p < .001$.

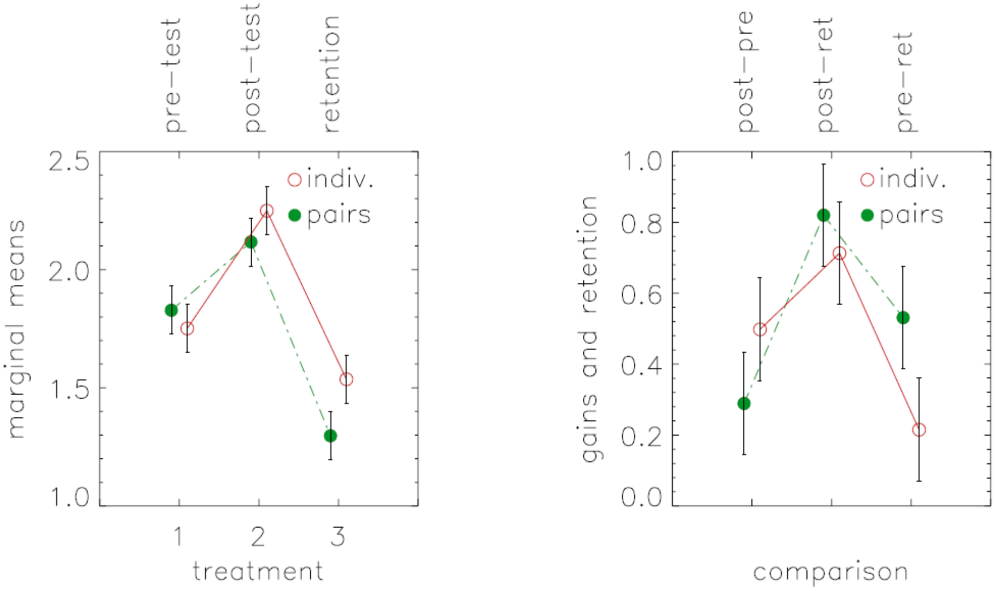


Figure 4.24: The left hand panel shows the test scores of the two groups of students, and the right hand panel shows the gains and losses between the various tests. It can be seen that students working as individuals outperformed the students working with peers by every measure. However, none of these measures was statistically significant at the p < .05 level.

Were there differences in gains between the groups?

As noticed in research question 3.1 and 3.2, individuals outperformed their peers working with peers. Again, the differences in gains did not breach the $p = 0.05$ significance level. *Individuals* ($M = .499$, $SD = .997$) and *Peers* ($M = .290$, $SD = 1.111$) groups ($t\ (197) = 1.459$, $p = .165$, $p > .05$), (Table 4.37). This can be seen in figure 4.25 and 4.26. The distribution of gains on the post-test for both groups is shown in figure 4.25. While the sizeable overlap between the two distributions does not allow for the conclusion that individuals significantly outperform their paired peers, it is certainly a noticeable finding. The paired-samples *t*-test was performed to quantify the significance level of the difference in gains. This test puts the significance level at $p = 0.165$.

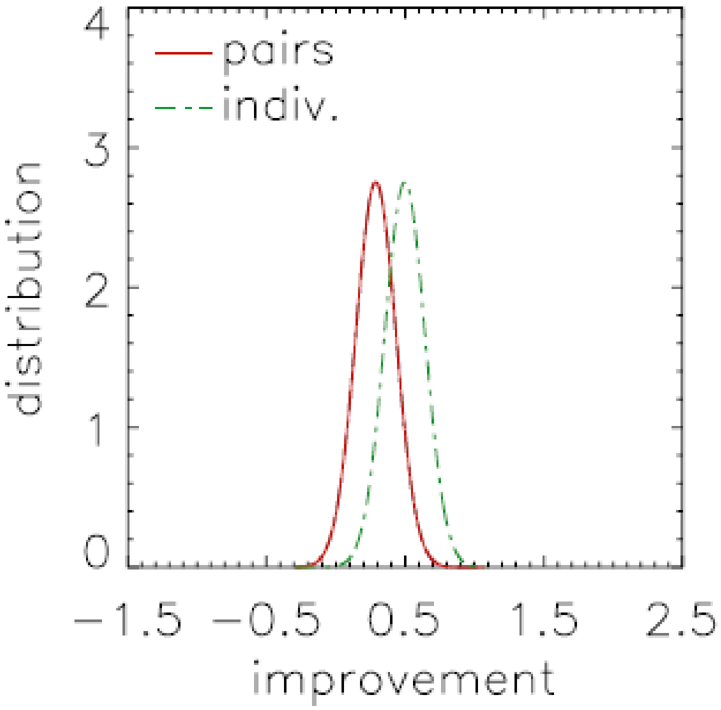


Figure 4.25: Individuals can be seen to outperform those students working with peers; however, there is a strong overlap between the two distributions so that the finding does not meet the $p = .05$ requirement.

Table 4.37: *The Paired-samples t-Test on Students Gains in the Individuals group and in the Peers group*

Group	N	Mean	SD	95% Confidence Interval for Mean	
				Lower	Upper
Individuals	99	.499	.997	.301	.696
Peers	100	.290	1.111	.071	.509

Note. **p* < .001, *p* = .165

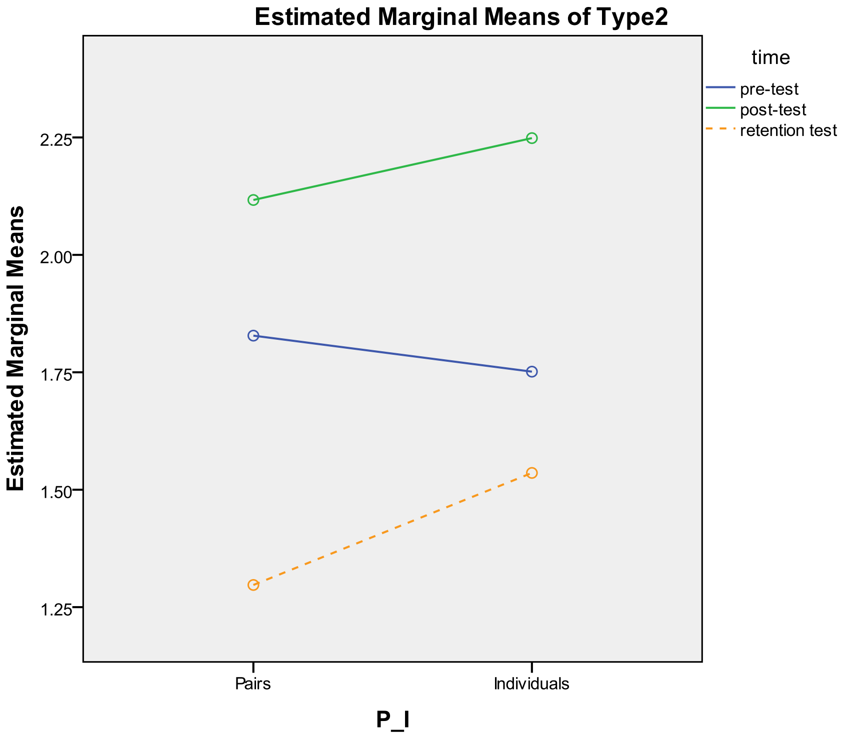


Figure 4.26: Clearly all students constructed new knowledge following the treatment, however, this knowledge was more than lost after 8-weeks (retention test).

Summary

Thus, it can be concluded that both groups of students were able to successfully transfer newly constructed knowledge of the concept to new situations regardless of whether they were working with peers or independently. However, also regardless of the way students constructed their knowledge, eight weeks later the gains have not been retained.

Overall Summary for Research Questions 3.1, 3.2 and 3.3

Surprisingly, students who completed the treatment on an individual basis outperformed their peers who worked in groups. While the difference in performance on the *Type I*, *Type II* or all the questions did not reach the $p < 0.05$ level, a clear tendency was observed in favor of individual work. However, these tendencies only pertain to immediate improvement following the treatment, long-term effects were not observed.

To explain the unexpected seeming advantage of working as an individual compared to working with peers the examination of the qualitative data was directed toward explaining differences between the individual learners and paired learners. These qualitative findings are presented in the next section.

III. Qualitative Investigation

The third research question of this academic investigation – whether students who work with peers construct greater knowledge of a scientific construct - was answered with interesting but inconclusive evidence from the quantitative analysis of the collected data. Hypotheses that students who worked in a computer-supported environment while collaborating with peers would construct greater knowledge and, as the result, score higher on *Type I* and *Type II* questions was not supported. Since there was not only (1) no statistically significant difference between the two groups of students in the expected direction but (2) students who worked independently consistently performed better than the peer groups even though it was statistically marginally significant. These unexpected findings were further investigated using qualitative inquiry.

The investigators were also interested in finding out how the participants of the study were constructing their knowledge of the concept through interactions with tutorials and peers and what new emerging themes for future investigations could be discovered. This section focuses on presenting the data of the qualitative part of the study with discussion of the findings in Chapter 5.

Thirty-six students out of 199 participants in the study were selected for interviews based on their individual performances on the post-test. Standardized gain was calculated for every participant and six students from each treatment group (learner- and program- controlled *SPIRUT*, *LTs*) where students who worked independently or were paired up with peers were selected for the interviews (Table. 4.39). A detailed description of the selection process is given on page 105 in Chapter 3. Also later in the chapter correlations between students' standardized gain and the nature of their interactions with tutorials and peers are presented in section *Peers and Individuals: Interaction and Performance*. Out of these students, 18 participants who worked independently and 18 participants who worked with peers were selected. Later in this chapter, the 18 students who worked with peers are referred to as the *Peer*s group and students who worked independently are referred to as the *Individuals* group.

Table 4.38: *Number of interviewed Peers and Individuals from each treatment group.*

Group	**LTs**	**SPIRUT**		**Total**
		Learner-controlled	**Program-controlled**	
Peers	6	6	6	18
Individuals	6	6	6	18
Total	12	12	12	

The interviews were semi-structured so that participants' answers could be compared during qualitative analysis. Students were asked several sets of questions: their prior familiarity with astronomy, use of computers and learning style, understanding of a concept, personal experience in working with a tutorial, and their experiences in working with a peer or independently.

Among other questions, students were asked to explain the stellar parallax concept verbally and to illustrate it by sketching the geometry of the concept (a stellar parallax model) on a piece of paper. In addition, students were asked to demonstrate conceptually and mathematically how they would find distances to stars by observing their apparent shifts. Participants were also asked to rate their experiences on a Likert scale about working with tutorials and peers and assessing how helpful the exercises were to them in learning the concept. Interviews lasted approximately 20-25 minutes on average. All interviews were audio recorded and transcribed for further qualitative investigation. Two co-raters created a coding schema, independently coded the interviews and then compared the assigned codes and negotiated meaning of transcribed students' words and sentences. The detailed process of the co-raters work is described on page 107 in Chapter 3.

The description of the qualitative part of the study is subdivided into the general scope and findings of the study, followed by some more detailed findings and by a part pertinent to the peers-versus-individuals difference in learning findings described in the quantitative part of the study. This section is concluded by findings that point toward emerging themes and directions for future research.

III.I. *Peers* and *Individuals*: Interviews

The interviewees' responses to the same repeatedly asked questions are shown in figures 4.27 - 4.43 and a full list of questions is included in Appendix D. All responses have been grouped according to whether students worked with peers or individually.

III.I.I. About the respondents

First, four ice-breaking- questions were asked to find out whether students had taken an astronomy course before, whether they wanted or needed to take the current course as their general education requirement. Students were also asked to self-assess their level of computer proficiency (see Appendix D for a detailed list of the questions).

Seven out of 36 interviewed participants had taken astronomy before in their high schools either as a separate course or a part of a science course: two were from the *Peers* group and five were from the *Individuals* group, (*About Responders,* Fig. 4.27). However, none of the participants could recall learning the stellar parallax concept before. As such, the group of interviewed participants constituted a good test-bed for testing the efficacy of the treatments.

Only eight students (five from *Peers* and three from *Individuals*) had an interest in astronomy and wanted to take an astronomy course in college while the rest of the participants had enrolled in the course as a part of the general science education requirement, (figure 4.27).

All 36 participants reported using computers every day and thirty-one participants described themselves as proficient users (16 from *Peers* and 15 from *Individuals*). Only five students (two from *Peers* and three from *Individuals*) were not very confident in their computer skills and said that they mostly use computers daily for checking their emails and Facebook. Follow up questions revealed that none of those students who described themselves as somewhat weak on this front felt intimidated by the *SPIRUT* environment. As such, the distribution of students over computer-based versus paper-based tutorials did not identify any complications for the use of computer-based learning as it can be seen in figure 4.27, *About Responders*.

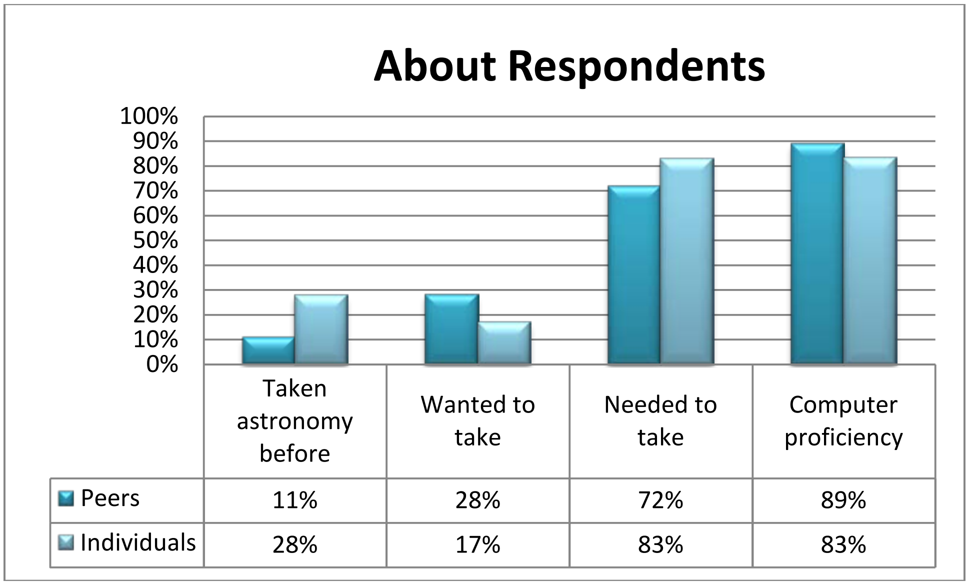


Figure 4.27: The column chart shows percent of respondents divided into subgroups of *peer*s and individuals. This chart visually demonstrates that there were no outstanding differences in the make-up of the groups when plotted versus their group memberships.

Understanding the stellar parallax concept

During the next set of questions, students were asked to provide a clear explanation of the concept of the stellar parallax and how it is used in astronomy. While 24 students (13 from *Peers* and 11 from *Individuals* groups) provided good explanations, seven students (three from *Peers* and four from *Individuals*) gave somewhat vague but acceptable answers. Five students (two from *Peers* and three from *Individuals*) provided incorrect answers confusing stellar parallax with Doppler Effect and Kepler's third law or they simply could not recall the concept. No outstanding difference between *Individuals* and *Peers* was observed, although students who worked with their peers did slightly better on this question, (*Understanding of the Concept*, Fig. 4.28).

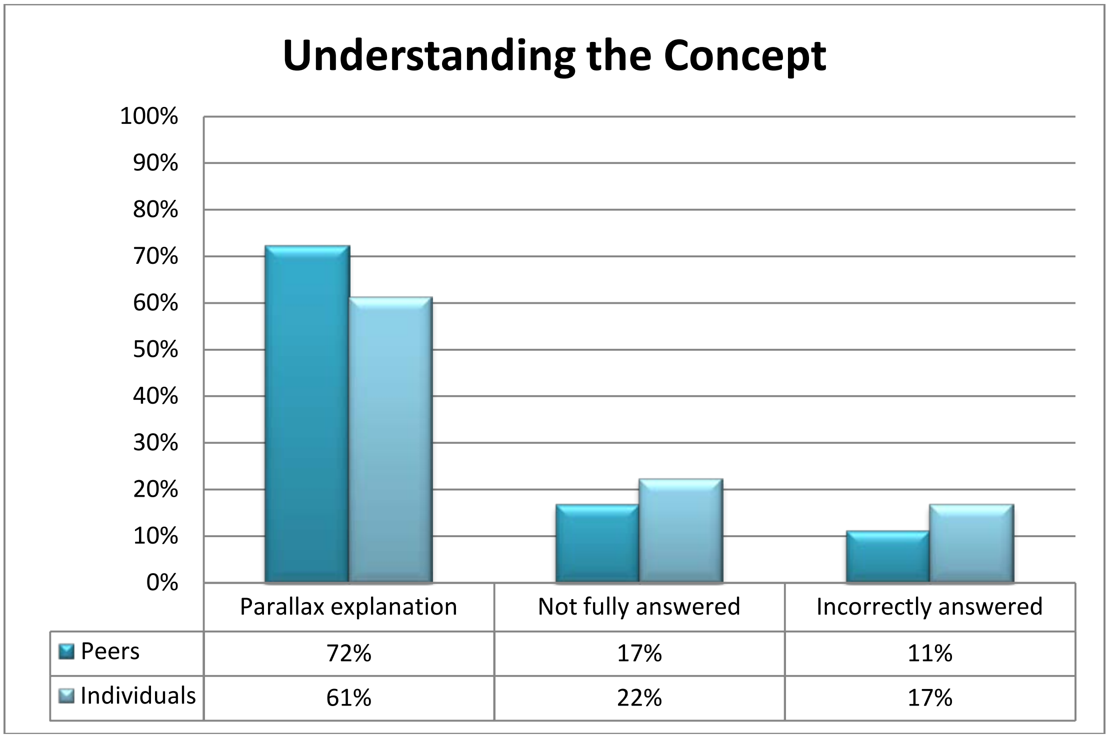


Figure 4.28: Most of the respondents from both groups were familiar with definition and did not have difficulties explaining the concept of the stellar parallax.

More than 20 students from both groups mentioned in their descriptions that stellar parallax is used to accurately determine distances to nearby stars (16 from *Peers* and 14 from *Individuals*) using the location of Earth in its orbit (12 from *Peers* and 12 from *Individuals*) six months apart (12 from *Peers* and 16 from *Individuals*) to observe the apparent shift of the stars (13 from *Peers* and 14 from *Individuals*) as they shift on the background of more distant stars. Overall, no outstanding differences were observed between those working with partners and those working on their own (*Stellar Parallax Components*, Fig. 4.29). However, interviewees from the *Peers* group tended to recall the distance component of the concept more frequently than other components of the stellar parallax.

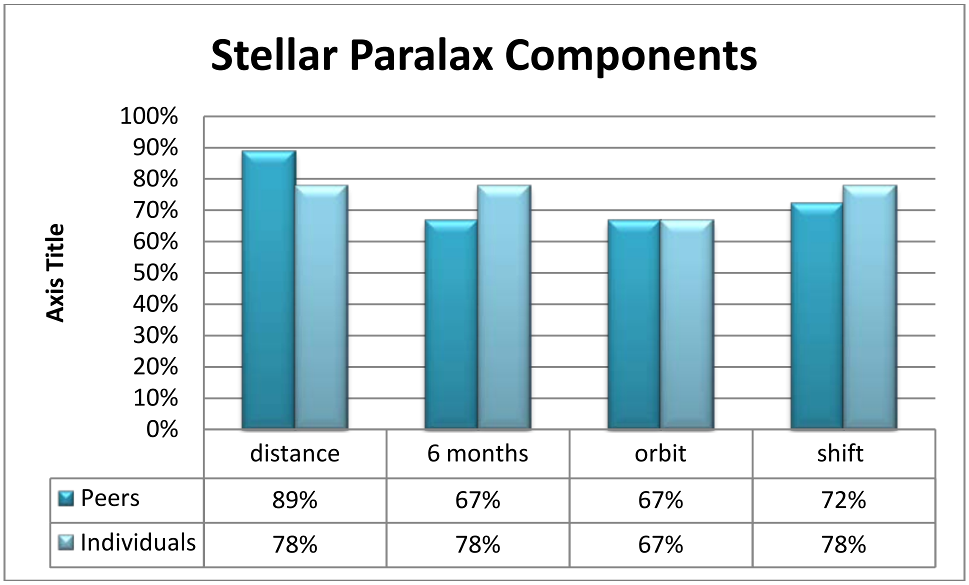


Figure 4.29: Questions relating to understanding the stellar parallax concept met with similar responses when subdivided into pairs and individuals.

To verify that students indeed had a factual knowledge of the concept, a practical question was asked later during the interviews -- students were given a hypothetical situation in which they would observe the apparent shifts of two stars over time. In this situation, one star would shift less than another. Fifteen students from the *Peers* group and 16 from the *Individuals* group correctly identified a star that shifts the most as the star closest to Earth (*Two Stars*, Fig. 4.30) therefore confirming that they were able to transfer their constructed knowledge of the concept to a simple practical application.

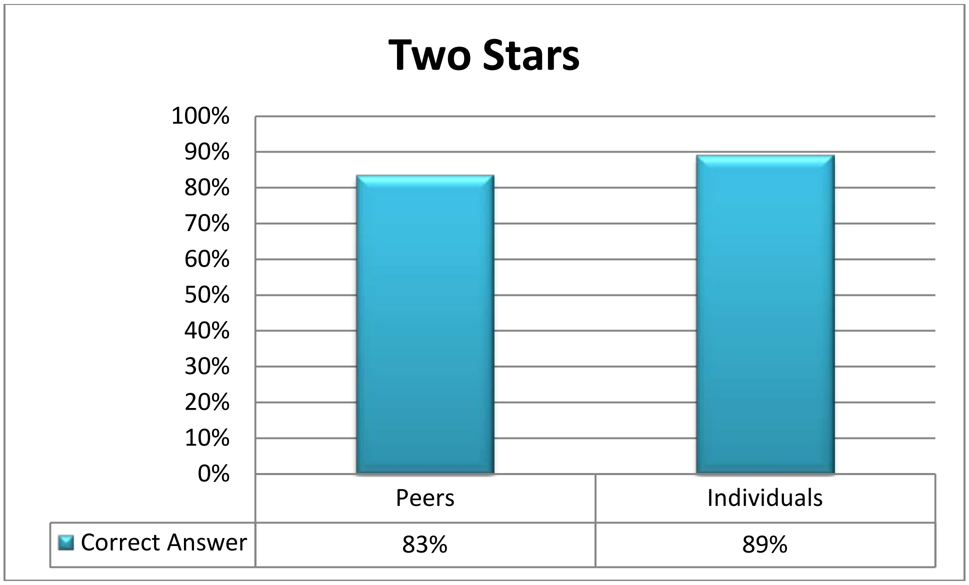


Figure 4.30: Questions relating to understanding the stellar parallax concept met with similar responses when subdivided into *Peers* and *Individuals*.

However, when students were asked to mathematically represent the concept, only five students from either of the groups could correctly recall the formula. Those ten students, who were either verbalizing or writing down the formula, were also asked what the components of the formula stand for. Only two out five *Peers* participants could correctly define the three formula components (distance, one astronomical unit, and parallax angle). However, there were no other students from the same group who could recall some components even if they did not remember the entire formula. At the same time four out of five *Individuals* participants gave accurate definitions of the formula components. In addition to these students, three more *Individuals* could correctly recall some of the components (*Recollection of the Formula*, Fig. 4.31). Overall, there was an advantage for those who worked as individuals in being able to explain what the equation stood for when compared to their peers who worked with peers.

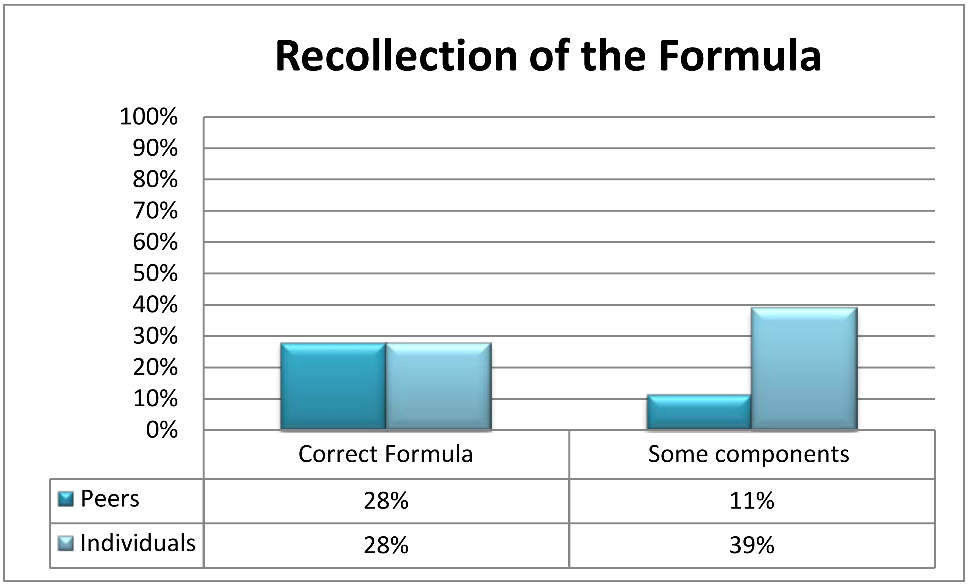


Figure 4.31: Only two students from the *Peers* group as compared to total of seven students from the *Individuals* group named and correctly explained all components of the formula.

It should be noted that during the interviews twelve students identified themselves as visual learners. However, six out of the 12 respondents who worked with paper-based tutorials (*LTs*) could not recall any visual representations of the concept. The other six interviewees had vague recollections of an exercise where they had to draw lines but they could not recall any details and final product of the exercise. At the same time, six students who worked with the interactive computer-based tutorial immediately recalled an animated figure and an interactive exercise that involved a ship in the sea (*Visual Learners*, Fig. 4.32). As further investigation revealed, most vivid students' memories of the stellar parallax concept were built around its visual explanation as an apparent shift of a ship on background of the distant stars. Another interesting aspect of this visual memory is that more students who worked with peers recalled images (eight students) than students who worked individually (four students) as demonstrated in figure 4.32.

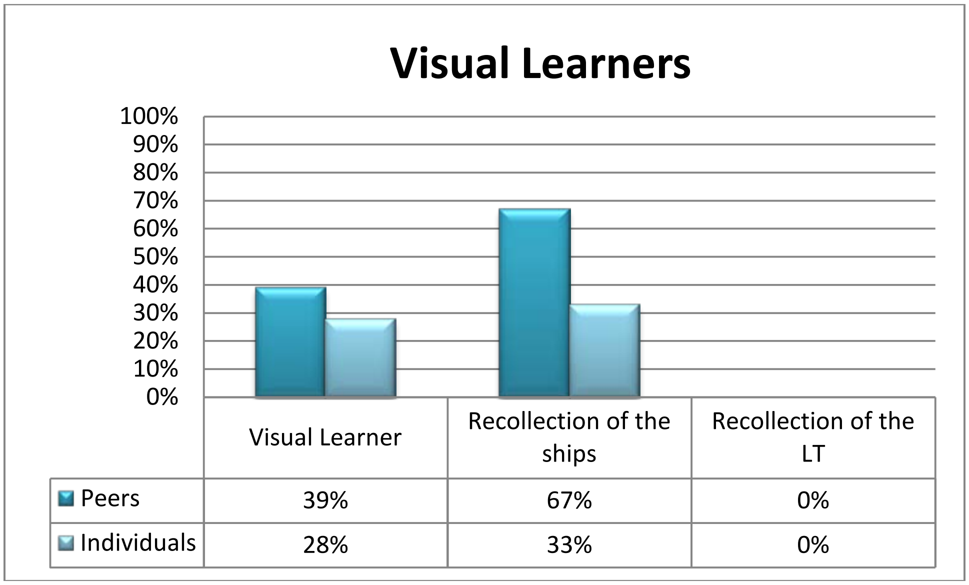


Figure 4.32: Students who worked with peers in *SPIRUT* environment had better recollection of visual representations related to the stellar parallax concept. Not one out of 12 students who worked with paper-based tutorials could clearly recall images and what lines they had to connect during their exercises.

While giving a description of the stellar parallax, students were asked to schematically represent the concept on a piece of paper. Only 15 out of 36 students provided accurate drawings (seven from *Peers* and eight from *Individuals*) (*Stellar Parallax Sketch*, Fig. 4.33). Some students rendered inaccurate sketches on which sides of Earth were used as two points of observations (*Students' Sketches*, Fig. 4.34a); some students represented their vague recollections of a base of a triangle with two points but could not represent how these points were related to an apparent shift (Fig. 4.34b). *Peers* and *Individuals* were equally challenged in performing this exercise.

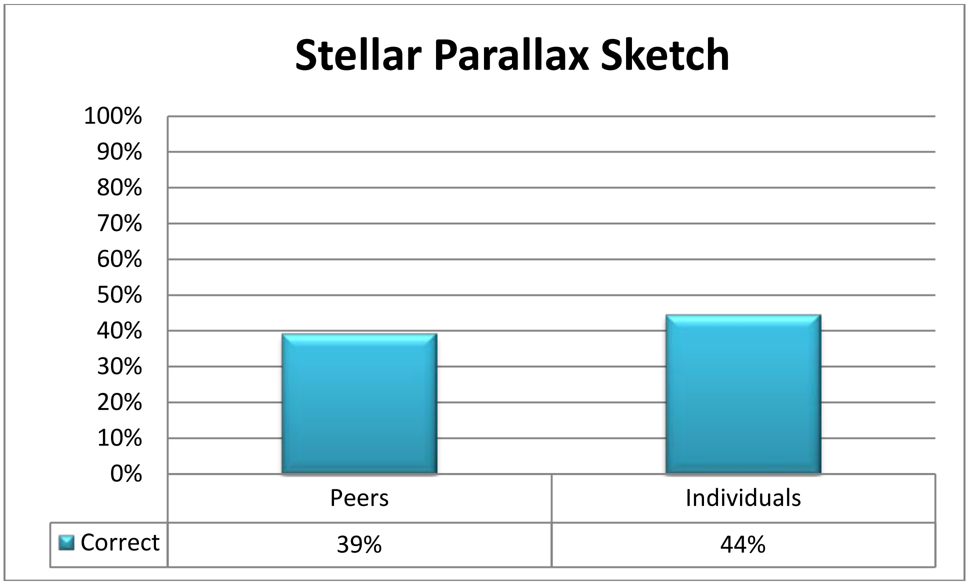


Figure 4.33: Less than a half of students who worked in either *Peers* or *Individuals* group could accurately sketch the stellar parallax model.

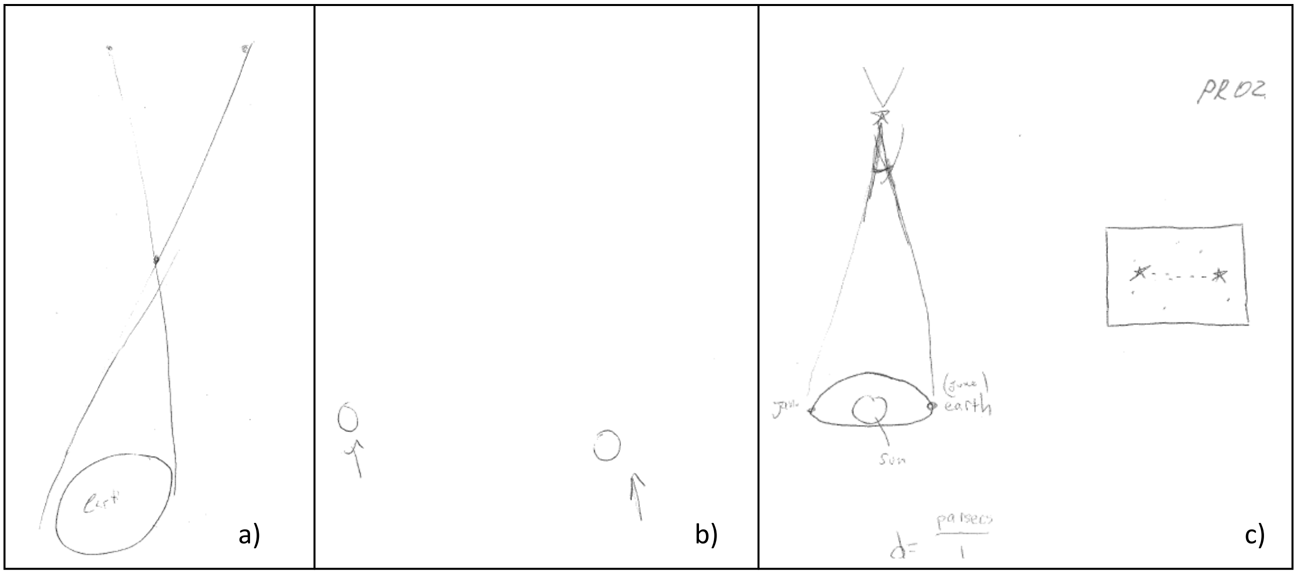


Figure 4.34: These sketches were drawn by students while they were explaining the concept of the stellar parallax: a) the students remembers a circle but does not know what its purposed in the concept just demonstrating a memorization of facts and a good understanding of the concept; b) this sketch demonstrates that the student has a vague memory that the parallax deal with a shift; c) this sketch shows that the students has a good grasp of the concept.

Although, more students from the *SPIRUT* group remembered animations and exercises than students from the *LTs* group, these visual memories were not reflected in the students' sketches of the stellar parallax. As is shown in figure 4.35, out of 15 students who correctly represented the concept on the paper, 10 students were from *SPIRUT* and five students were from *LTs* (*Sketch: SPIRUT vs. LTs*, Fig. 4.35). While *Peers* from the *LTs* group had better understanding of the concept, *Individuals* from the *SPIRUT* group outperformed students who worked on the tutorial with their peers.

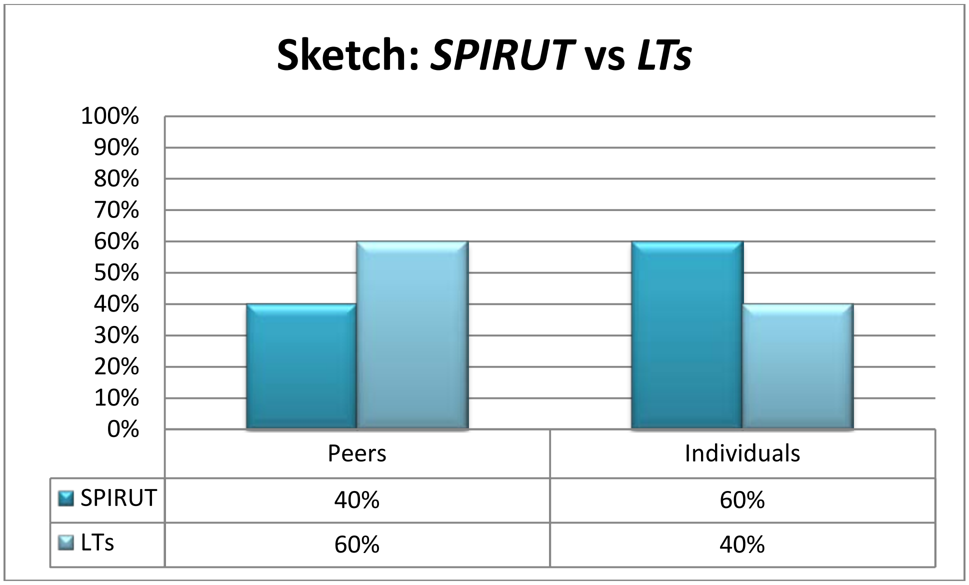


Figure 4.35: This figure shows distribution of 15 students (10 from *SPIRUT* and five from *LTs*) who correctly sketched the concept.

To find out what was the source of information that had the most influence on participants' memories of the concept, the interviewees were asked what was most helpful to them in learning the concept: the lecture preceding the treatment, printouts from the course' textbook, or exercises with tutorials. Interestingly, eleven students who worked with peers considered the 20 minute lecture to be the most helpful, while eight students who worked individually considered the tutorial to be almost as helpful as the lecture, (*Source of Information*, Fig. 4.36).

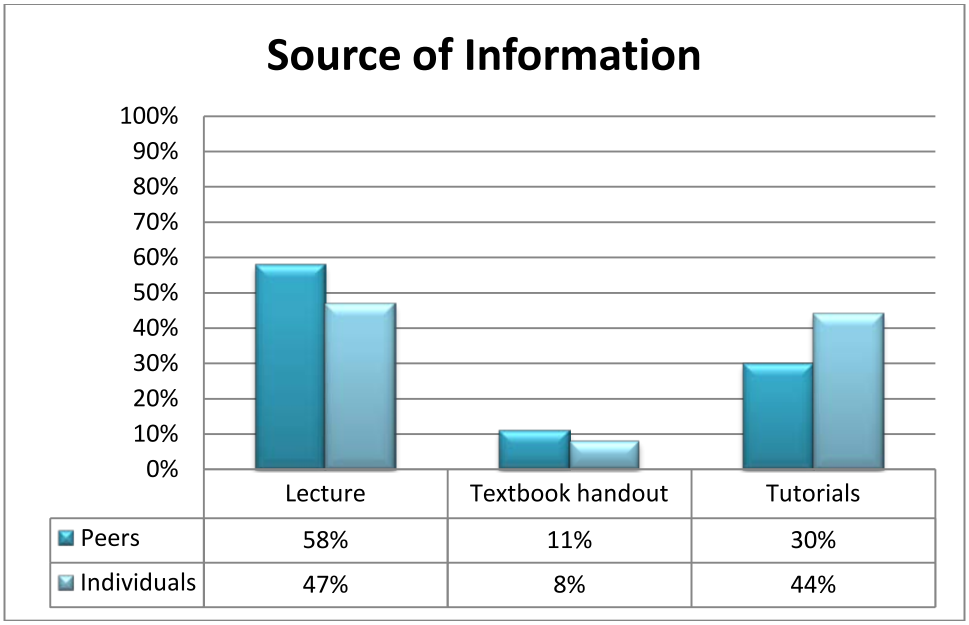


Figure 4.36: From *Peers,* 11 students named the lecture as the primary source of recalled information, 2 named the textbook handout and 5 named tutorials. From *Individuals*, 8 students named the lecture, 2 – the textbook printout, and 8 – tutorials.

Based on the students' responses in this part of the interviews, the participants had greater conceptual understanding of the stellar parallax as they had a grasp on the concept's mathematical representation. Overall, there were no substantial differences found between the two groups of the respondents accept for *Individuals* recalling more components of the formula and attributing their constructed knowledge equally to tutorials as to lectures.

Reflecting on the tutorials

This set of questions was asked only of *SPIRUT* users. The questions aimed to elicit the nature of participants' interactions with the tutorial and each other in order to discover whether the differences in performances between *Peers* and *Individuals* lay among this group of students. Twenty-four students, 12 *Peers* and 12 *Individuals* were asked questions about their experience working with the program and with their partners.

During the interviews three students from the *Peers* group mentioned that some questions and exercises from *SPIRUT* were hard and two more students thought that some parts were confusing. At the same time, five students who worked independently shared that they were intimidated by the hard beginning of the tutorial that assumed that they already understood the concept and could perform the required calculations in order to determine a distance to a star. Two out five students and two additional students from *Individuals* also were occasionally confused while working on the exercises. Overall, *Individuals* made these comments at about twice the rate of students who worked with peers (five from *Peers* and nine from *Individuals*), (*Hard and Confusing Tutorials*, Fig. 4.37).

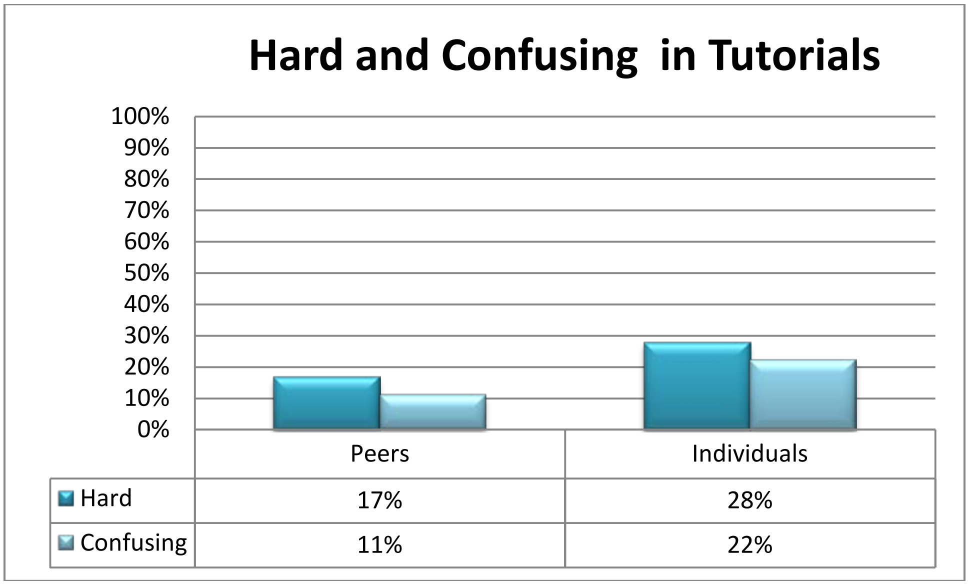


Figure 4.37: This chart shows that while only 5 students from *Peers* and 9 from *Individuals* out of 36 respondents thought that *SPIRUT* was hard and confusing at a time, students who worked independently mentioned this fact almost twice as often.

At the same time, 17 out of 24 interviewed students from the *SPIRUT* group mentioned helpfulness of feedback as an immediate response to their answers on multiple-choice questions (seven form *Peers* and 10 from *Individuals*). Students who worked independently found the feedback being especially helpful to them in knowing how well they were performing on the tasks, (*Feedback in SPIRUT*, Fig. 4.38). In addition, all 12 interviewed students from the *LTs* group shared that lack of a feedback kept them lost and guessing since they did not know whether their answers were correct.

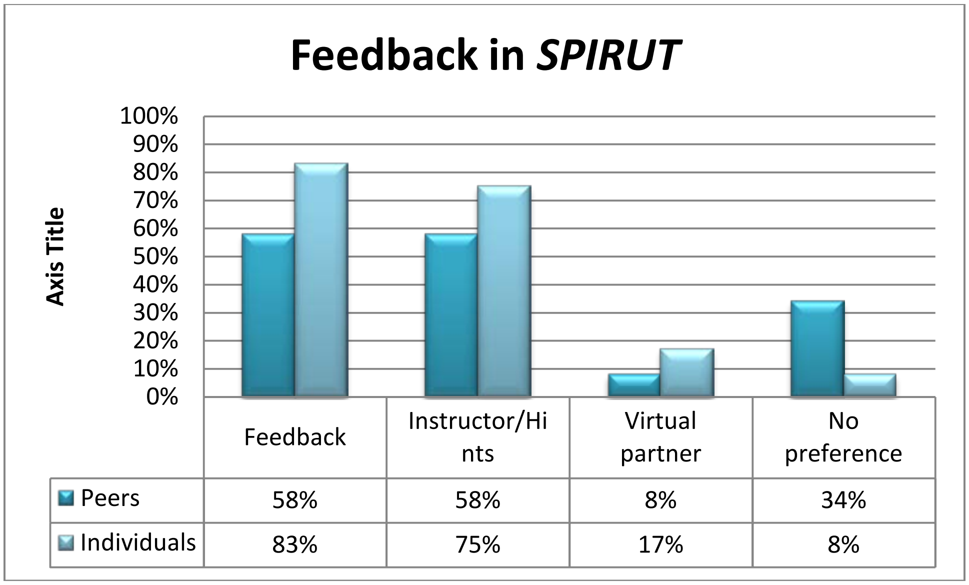
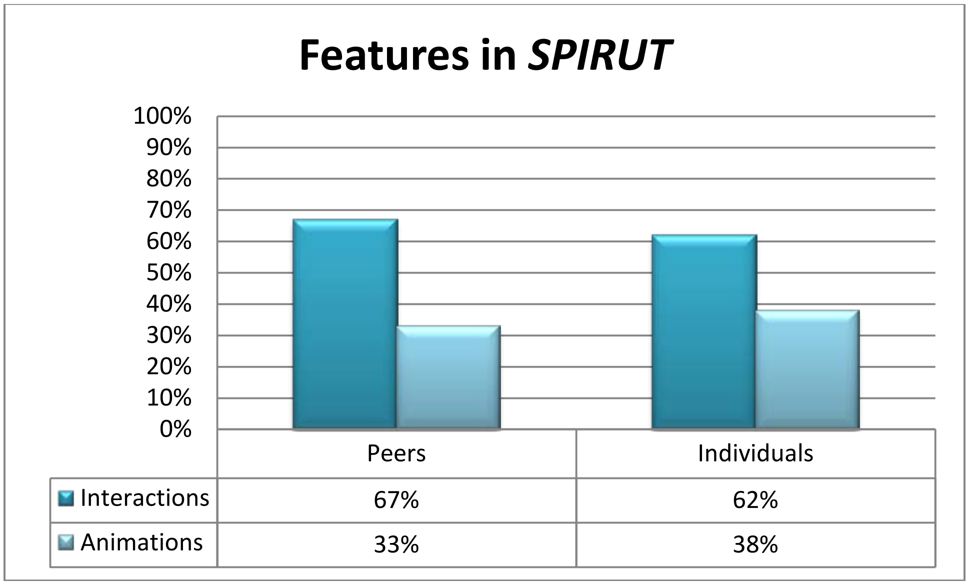


Figure 4.38: This chart shows that students who worked independently expressed a stronger need to have a feedback and preferably as a hint rather than having a virtual partner.

As a follow-up question to students' comments on feedback in *SPIRUT*, the students were asked what form of feedback they prefer: to receive feedback from the tutorial reacting to their answers as an instructor and hint them toward a right answer, or rather be a virtual partner and help in forming of a second opinion or another idea. Sixteen students wanted to have hints versus a virtual learning partner, with individuals perhaps slightly more keen on direct hints (seven from *Peers* and nine from *Individuals* groups). Only three students wanted to see feedback in a form a virtual partner providing a second opinion or asking additional questions as a form of feedback (one from *Peers* and two from *Individuals*) and remaining five students did not have a preference (four from *Peers* and one from *Individuals*).

Twenty-four students who worked with the computer-based tutorial were asked whether *SPIRUT*'s interactive exercises or animations had the most profound impact on the way they learned the concept. Most of the students thought that while animations often clearly showed them how things worked, their interactions with the tutorial helped them the most. *Individuals* and *Peers* gave similar responses. Eight *Peers* and seven *Individuals* thought that using tools to adjust scenes and observe the changes were most helpful to them. Four *Peers* and four *Individuals* thought that animations were very helpful to them, (*Features in SPIRUT*, Fig.4.39). One student who worked independently could not decide what was the most helpful to her.



	Peers	Individuals
Interactions	67%	62%
Animations	33%	38%

Figure 4.39: This chart shows that 15 out 24 students preferred interactions to animations as a learning tool. Only 8 out of 24 students preferred animations to interactions.

Students who worked with *SPIRUT* were asked what types of questions were most helpful to them in learning the concept: multiple-choice or essay type questions. A clear difference emerged between *Peers* and *Individuals*. Eight students who worked independently preferred multiple-choice questions and only three would rather answer essay questions while one student saw benefits from both types of questions. Students who worked with partners did not have such a clear division: three preferred multiple choice questions and five preferred the essay questions, and the remaining four students did not have a strong opinion, (*Type of Questions*, Fig. 4.40).

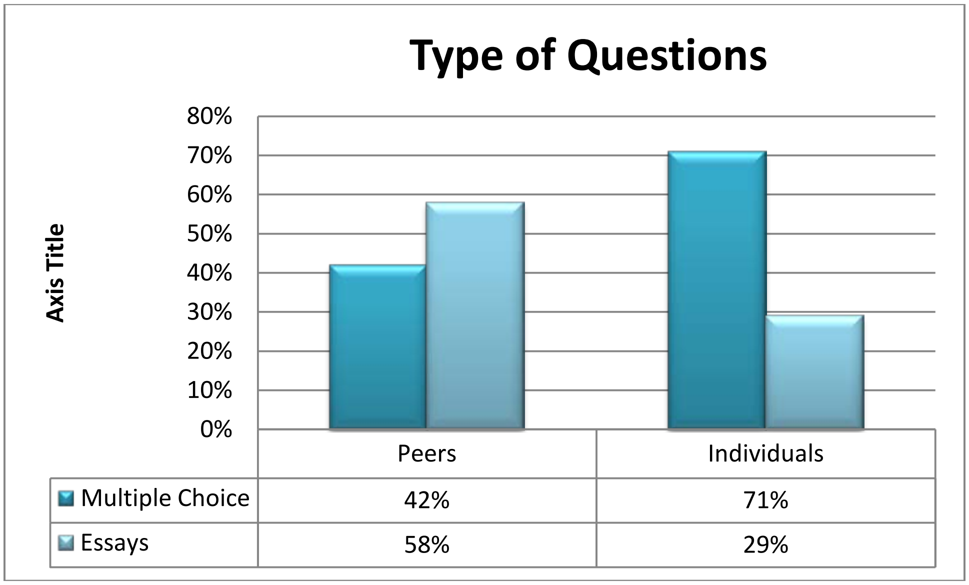


Figure 4.40: This chart demonstrates that more Individuals preferred multiple-choice to essay questions.

Another question asked students about their experience working in a learner-controlled or a program-controlled environment of *SPIRUT*. Students were asked to comment whether they liked to be in control of their actions or they would rather trust the tutorial and allow *SPIRUT* to redirect them to different parts of the tutorial based on the users' performance. Six out of 12 students who worked in the learner-controlled environment preferred that environment to the program-controlled version of *SPIRUT*. Four students said that they would rather work in the program-controlled *SPIRUT* and two students thought that it would not make any difference to them. Four out of 12 students who worked in the program-controlled *SPIRUT* preferred that version of the tutorial, three students wanted to try a learner-controlled, and five did not have a strong opinion about it. *Peers* students and *Individuals* gave similar responses with a slight preference of *Peers* to control their actions (*Type of SPIRUT*, Fig. 4.41).

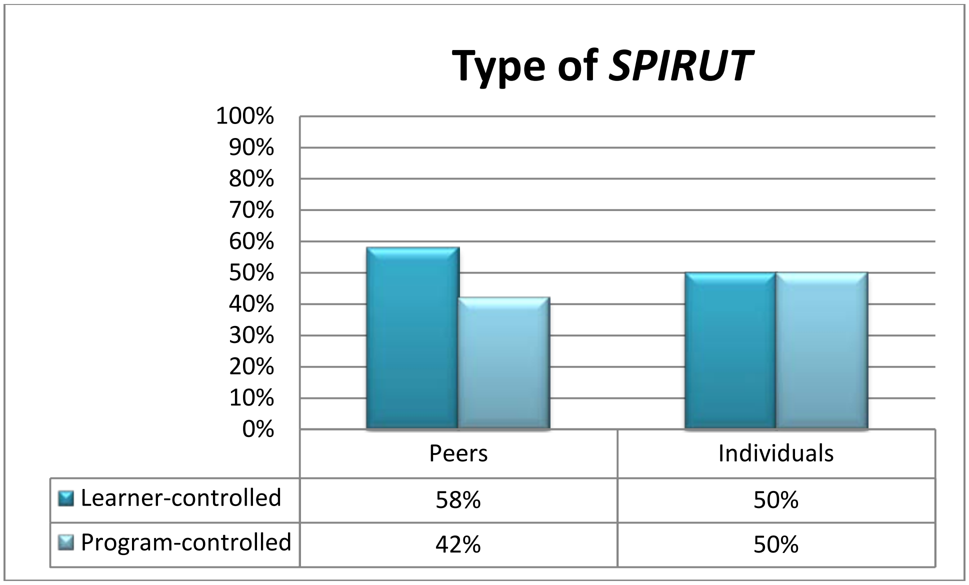
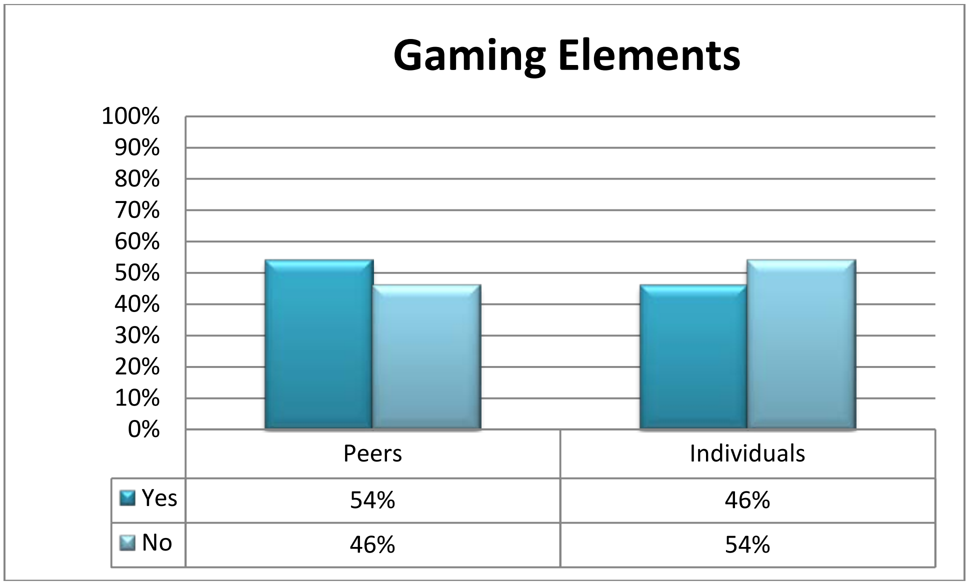


Figure 4.41: This chart demonstrates that *Peers* saw more benefits from being in control of navigation in the environment.

Additionally, students were asked whether they would be more motivated if *SPIRUT* included gaming elements, such as competition or earning a certain amount of points. While students' opinions split in this case, most of the students across all groups did not have a strong opinion. Only three interviewees strongly advocated including gaming elements in the tutorial and five thought that such addition would be distracting for leaning. The remaining 16 students thought that it may go either way, (*Gaming Elements*, Fig. 4.42).



	Peers	Individuals
Yes	54%	46%
No	46%	54%

Figure 4.42: This chart demonstrates that regardless of the way students interact with the environment, gaming elements is not an important component of a computer-based learning environment.

Based on students answers to the question in the part of the interviews, it was discovered that *Individuals* more frequently mentioned that the tutorial was hard, they appreciated feedback more (especially in a form of hints), and they preferred multiple-choice to essay questions. Peers, however, preferred the essay form of questions, and they liked the learner-controlled environment.

<u>*Working on exercises*</u>

In this group of questions students were asked about their experiences of working alone or with partners. Students who worked either independently or with partners were asked whether they would prefer to work as individuals or in a paired group with peers.

Six students who worked independently thought that they would learn more if they were working with a partner, six wanted to work independently and six did not have a strong opinion. Four students from *Peers* thought that they would not learn as much as they did if they were to work by themselves. Eight students liked and wanted to work with a partner and six participants did not have a strong opinion. (*Working Together*, Fig. 4.43).

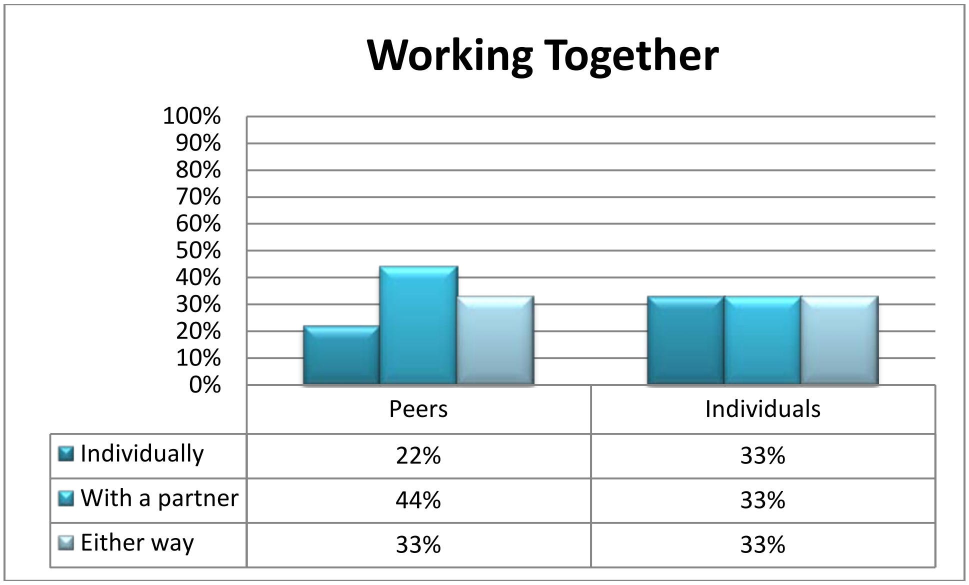
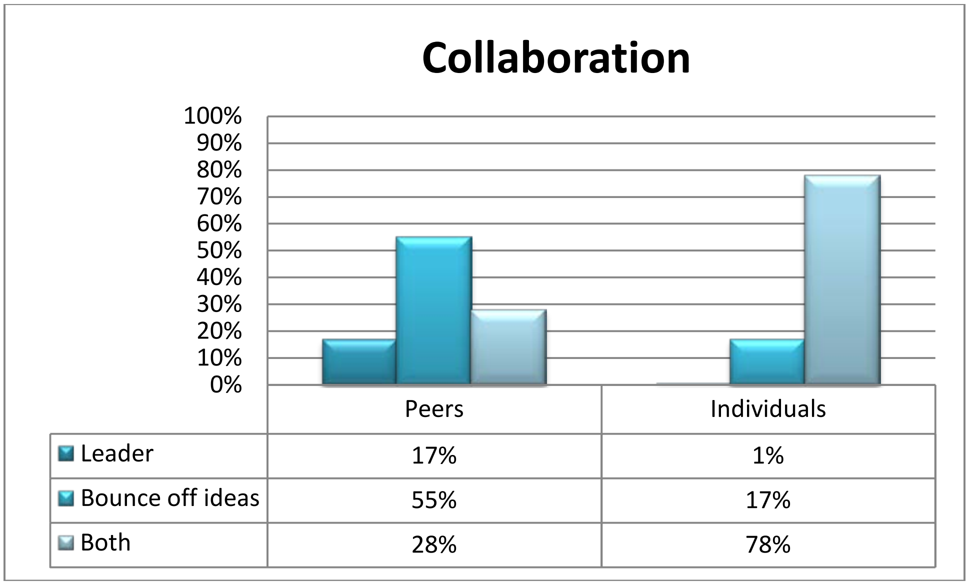


Figure 4.43: This chart shows that students from both groups did not have a very strong opinion on working with or without a partner.

In the next question, students who worked with peers were asked about the nature of their collaborations with their partners and students who worked independently were ask to suggest how they would work with a partner if they were assigned to work with one. Only three students from *Peers* identified themselves as leaders, ten students described their work as bouncing ideas off each other in more of a supportive role and five students thought sometimes they felt they were being leaders and sometimes were supporters, (*Collaboration*, Fig. 4.44). While fourteen students from the *Individuals* group who were asked to predict what type of collaboration could take place if they were to work with a partner, did not have a strong opinion and thought that any type of collaboration can take place, one student was certain of assuming a leader position and three thought they would be either in support of a leader or would equally share their ideas while working on the exercises.



	Peers	Individuals
■ Leader	17%	1%
■ Bounce off ideas	55%	17%
■ Both	28%	78%

Figure 4.44: This chart demonstrates that equally shared participation was most popular among *Peers*.

As a follow up to this question, students from the *Peers* group were asked to compare their level of knowledge of the stellar parallax to a level of knowledge of their partners. Six students thought that their understanding of the concept was superior to the one of their partners. Eight students thought that their understandings were equally matched and four students felt that their knowledge was inferior compared to their partners, (*Peers: Self-Assessment*, Fig. 4.45). It should be noted, however, that their self-assessment did not always match their scores on the post-test.

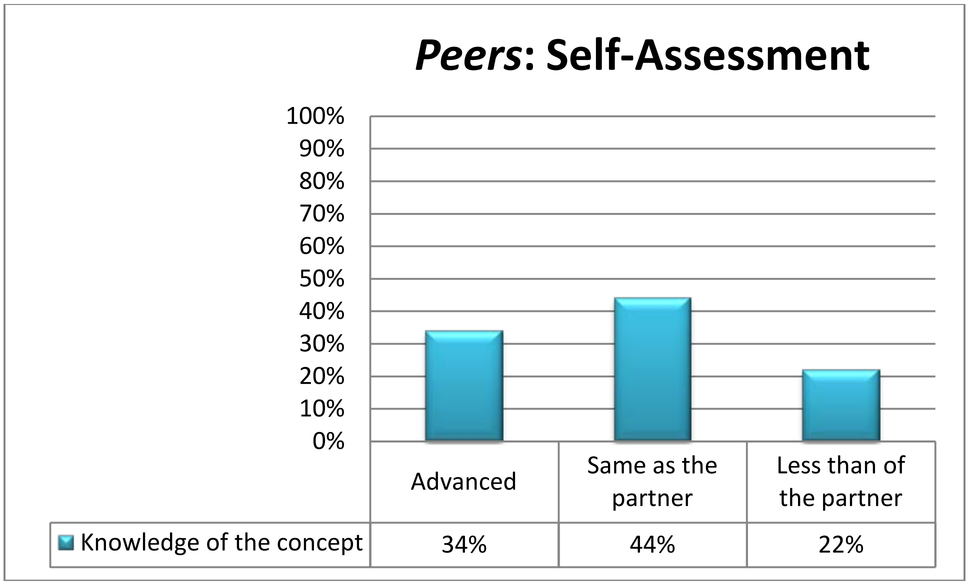


Figure 4.45: This chart shows that most of the students think about their knowledge of the concept either equal to or superior to their partners.

Twelve students who worked with the learner-controlled version of *SPIRUT* were asked to describe their work process with the tutorials. Nine out of 12 students who worked in the *Peers* group skipped answering questions on many occasions. Only two out of 12 students used their control over the tutorial to go back to review their previously answered questions and completed exercises, (*Workflow*, Fig. 4.46). Students from the *Individuals* group had a different pattern of behavior: six out of 12 students used their control over the tutorial to skip the exercises and six students proceeded in a step-by-step manner through the tutorial. Four out of 12 *Individuals* went back to check their answers.

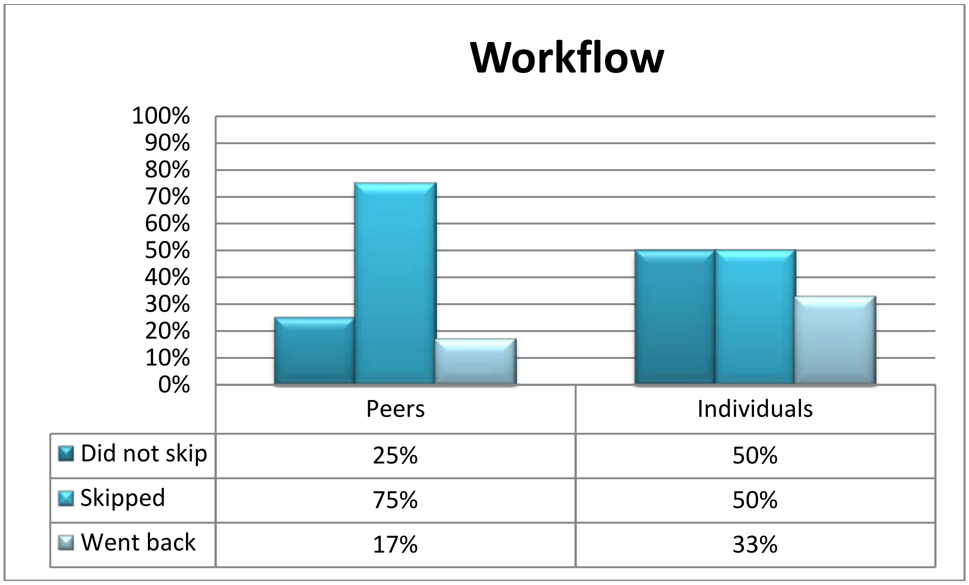


Figure 4.46: This chart shows that *Peers* skipped questions and exercises more frequently than *Individuals*.

As is seen from students' responses to questions in this part of the interviews, *Peers* preferred to work with partners in a form of bouncing ideas off of each other and they tended to skip exercises and questions to go forward in the tutorial. *Individuals*, however, used their control of the environment to go back to retrace their answers and they rarely skipped any questions and exercises.

III.I.II. *Peers* and *Individuals*: Interaction and Performance

Standardized Gain: Three Subgroups

For further analysis of the qualitative data, students' answers were scrutinized and compared against students' standardized gains on the post-test.

To interpret the qualitative data, the 36 interviewed students from both the *Peers* (*Peers*, Fig. 4.47) and the *Individuals* groups (*Individuals*, Fig. 4.48) were divided into three sub-groups based on their standardized gain (Table 4.39):

- A, high gain from .68 to 1 and medium gain from .34 to .67
- B, low gain from 0 to .33
- C, negative gain below 0

Table 4.39: *Division of the interviewees into three groups according to their level of standardized gain.*

Group	**A** **high/medium gain:** **(.68- 1.0)**	**B** **low gain:** **(0.0-.33)**	**C** **negative gain:** **(< 0.0)**	**Total**
Individuals	9	4	5	18
Peers	5	7	6	18
Total	14	11	11	36

Group A (14 students, nine from the *Individuals* group and five from the *Peers* group) can be described as two types of students: those who did not have sufficient knowledge of the concept before working with tutorials and thus benefited the most from interaction with the tutorials with and without peer assistance, and those students who had already had a good grasp on the concept and improved their understanding of the concept by doing exercises either individually or with peers (Figs. 4.47 and 4.48).

Thirteen out of 14 students in this group had never taken an astronomy course before and were taking the course as a fulfillment of their general education requirement. They selected the course because they had a mild interest in astronomy and chose the course over taking physics or biology courses. During the interviews, eight respondents shared that working with tutorials helped them to understand the concept better, three were recalling what they knew about the parallax from the given lecture, and one student was not sure whether the lecture, a passage from the textbook, or the tutorials helped him the most. It should be mentioned that, when asked, eleven students were able to provide a detailed explanation of the concept while just one of them could only recall that the stellar parallax is a method of finding distances to the stars.

At the same time only eight students were able to mathematically represent the concept ($d = 1/p$) and seven of them could explain each component in the formula. All 12 students had no doubts that if they observed two stars that appeared to shift a different amount over time, the one that shifts the most is closer to Earth than another star. Eight students in this group preferred working with peers, two would rather work individually and two did not have any preferences.

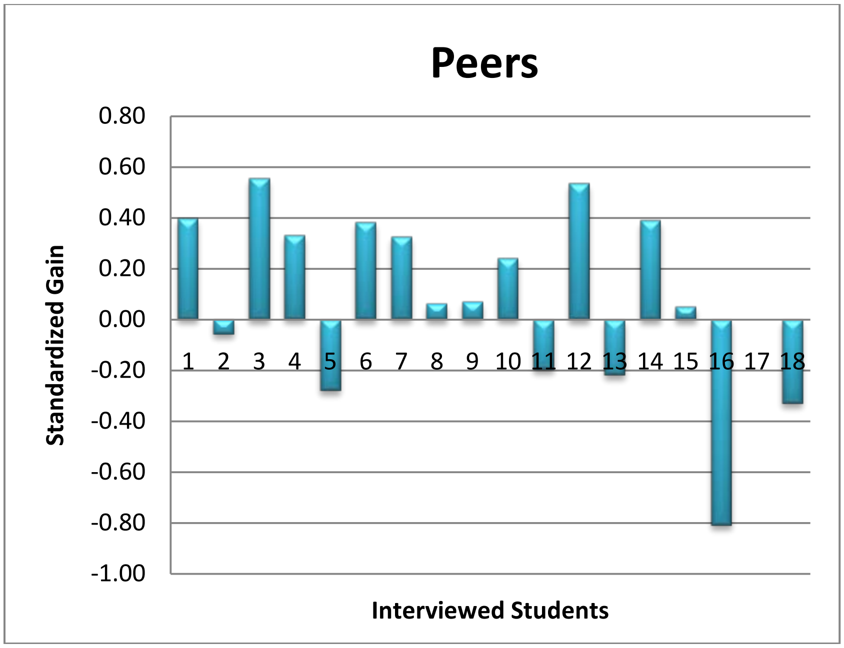


Figure 4.47: The distribution of standardized gains of the *Peers* group of the interviewees, subdivided over their grouping during the treatments.

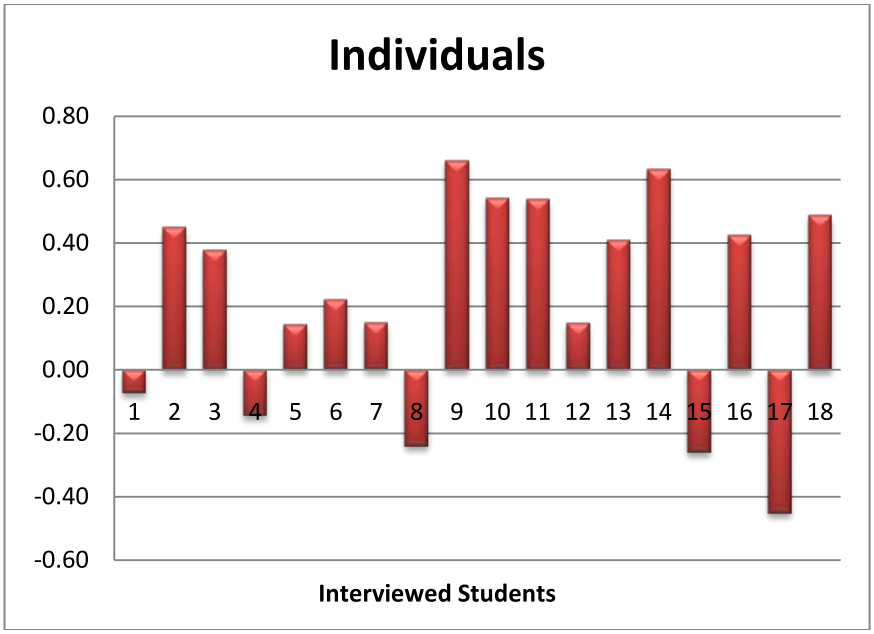


Figure 4.48: The distribution of standardized gains of the *Individuals* group of the interviewees, subdivided over their grouping during the treatments.

Group B included eleven students (four from the *Individuals* group and seven from the *Peers* group) who either did not construct any knowledge after working with tutorials or had only modest improvement (up to .33 points gain) as a result of already having a good understanding of the concept.

Two students from the *Individuals* group in this category had below average scores on the pre-test (*Individuals*, 5 and 12, had scores of 5.9 and 3.15 respectively) and did not show significant improvement on the post-test (7.06 and 4.75 respectively). Two other students had higher scores on the pre-test (*Individuals*, 6 had a score of 7.93 and 7 had 6.84). On the post-test, student number 6 received a high score of 9.27, while student number 7 performed below average and received 7.9 points, (Fig. 4.48).

Among the students who worked with peers and fell in the low gain category three of them showed gain between .24 and .33 and performed below the mean on the pre-test (*Peers*, 4: 5.73, *Peers*, 7: 7.11, and *Peers*, 10: 6.15) and performed much better on the post-test (8.48, 9.35, and 8.05 respectively.) Two out of four remaining students had high scores on the pre-test (*Peers*, 8: 8.21 and *Peers*, 9: 11.56) and only showed modest improvement (8.57 and 11.73 respectively) while two other students did not show gains after working with tutorials (*Peers*, 15: 5.31 and 5.73; *Peers*, 17, 6.85 and 6.86.)

Group C includes eleven students (five from the *Individuals* group and six from the *Peers* group) who instead of constructing better knowledge after working with tutorials showed a decline in their original understanding of the concept that they have constructed after the presented lecture and reading a handout. Their standardized gain was in the range of -.06 to -.81 points. Overall, their scores on the pre-test were 6.47 points which were .21 above the mean. Only one student in the group had a significantly higher score of 8.25 points and dropped down to 6.75 on the post-test (Figs. 4.47 and 4.48).

In this group only five out of 11 students gave a clear explanation of the stellar parallax concept and one student was able to recall the formula. Eight students mentioned that they learned the most about the concept from the 20-minute lecture. As compared to groups A & B, more students in this group preferred multiple-choice questions to essays: "Because sometimes I didn't really know what to put. So with the multiple-choice you could make an educated guess" (IU11), they did not have a strong preference of having any type of feedback, and mentioned being constrained by time. In addition, two out of six interviewed students from *Peers* shared that they would rather work individually, one student was a leader in her team and six students thought that working with peers could be beneficial for sharing ideas and negotiating answers.

Interactions within the Subgroups

Based on analyzed students' responses during the interviews and comparison with their scores on three assessment tests, the following patterns were revealed concerning students who worked in peer groups. In most cases, students with higher gains benefitted from working with a peer irrespective of the level of their partner, whereas students with lower gains experienced the opposite. Students who were well matched did better (with a peer) than mismatched students. In addition, learning appeared to have been made harder in some cases because of time constraints that either one of the students with peers could have, influencing the time the other student could spend on the tutorial with negative results in overall learning (*Peers: Standardized Gain*, Fig. 4.32).

Five groups of paired students showed gains for both partners (groups 1, 4, 5, 9, and 11). Groups 1, 9 and 11 were students from the A subgroup (high standardized gain) who had comparable results on the pre-test with their partners and shared leadership while working on the task (groups 1 and 9) or took the leadership (group 11). In groups 3, 5, 8 one of the partners gained while another one lost. Interviewed students from these groups had various pre-tests scores as compared to their partners but they all improved their scores as a result of sharing leadership during their work on the tasks while their partners not only did not benefit from the collaboration, but even lowered their scores. Groups 10 and 13 were from the C subgroup and worked with the paper-based tutorial. Two interviewed students from the groups originally had higher scores than their partners, and both students participated in shared leadership type collaboration. In group 15 both students had above average pre-test scores and shared leadership during their work with paper-based tutorials, however both students received lower scores on the post-test.

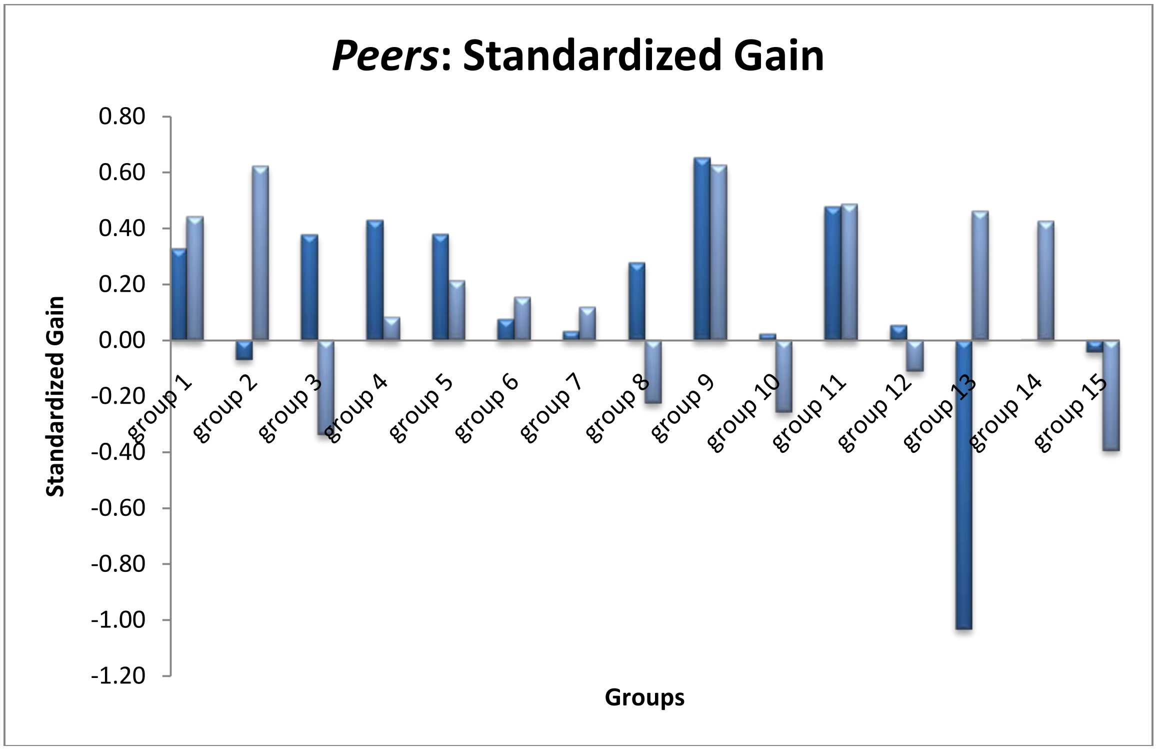


Figure 4.49: Of the 18 students interviewed who had worked in peers, their standardized gains were compared to their partner. Students in group 2, 3 and 8 were both interviewed, of the remaining students their partners were not interviewed.

Whether students who worked with peers constructed greater conceptual knowledge or lessened their understanding of the concept can be partially explained by the nature of the students' collaboration (or lack thereof) and how they perceived their role in it. Students in subgroup B saw themselves either as equal partners and appreciated working with a partner as a result of ability to compare ideas or they did not have strong preferences. According to students’ answers from subgroup C, they preferred working with partners so they could learn better from sharing ideas (Table 4.40).

Table 4.40: *Table the Peers group divided three subgroups according to their standardized gain and preferences on working with or without a partner*

Peers				
Wanted to work in:	**A**	**B**	**C**	**Gain per member**
Peers	3 [L, B, N]	1 [B]	4 [B, B, B, N]	0.09
Individually	1 [L]	2 [L, N]	1 [B]	-0.03
No preference	1 [B]	4 [B, B, N, N]	1 [B]	0.06

NOTE: students in pairs saw themselves as: *L*– leaders, *B* – bounce off ideas, *N* – had no preferences

Students who worked individually and preferred to work that way had the highest gain per group members. Individuals who would rather work with a partner but were randomly assigned to work solo still had a noticeable gain. Students who did not have a preference on how to work on tutorials had the least gain per group members (Table 4.41).

Table 4.41: *Table of the Individuals group divided into three subgroups according to their standardized gain and preferences for working with or without a partner*

Individuals				
Wanted to work in:	**A**	**B**	**C**	**Gain per member**
Peers	2	3	1	0.24
Individually	5	0	1	0.38
No preference	2	1	3	0.05

NOTE: The table was calculated using a sum of standardized gain in the row divided by the number of member in the row.

III.II. New and Emerging Themes

The responses from the interviews were also examined to identify new themes about how students work and learn from tutorials. For this purpose, three approaches were used to examine the interview responses: Pawing, Word Frequency Count, and Cutting and Sorting. During the pawing process color-coded and highlighted words and sentences in the students interviews was used to identify similarities between answers, what students like or did not like, and what they spontaneously brought up during the interviews. The word frequency process helped to identify what words were mentioned the most by students from all six groups. Thus it was noticed, that students in the *LTs* group did not use word like satisfy and exciting while students from *SPIRUT* used these words 17% of the time. At the same time, students who worked in the *LTs* group mentioned the word frustrating only .2%, while students in the *SPIRUT* group mentioned it 2.5%. While the reason for using these words is not known without the context, these differences in word frequencies led to further detailed study of the content.

Time constraints: regardless of whether students worked individually or with peers, they wanted to be able to work from home and to not have a time constraint. Many students pointed out that they had to rush through the exercises and post-test because they had other classes scheduled or a time for a favorite TV show was approaching.

Lack of knowing terminology: The interviews also revealed that students' use of terminology was very poor. They did not know the difference between rotation and revolution, they had a hard time recalling the word parsec and many of them replaced it with a made up words such as parsecond or paraseconds.

Most of the students failed in writing the formula ($d=1/p$) down. It should be pointed out that students who remembered the formula could also verbalize it by explaining each component in it. Those students who wrote the formula incorrectly could not explain what the "*p*", "*1*", and "*d*" meant in the equation. Many of the latter students thought that "d" stands for the distance between Earth and the Sun. This misconception was equally shared between students from *SPIRUT* and *LTs* groups.

Grade pressure and anxiety: many students from both groups thought that they would do better on the exercises if they were not afraid to provide a wrong answer and thus lower their grades.

Drawing exercises: nine out of 12 students (four from *LTs* and five from *SPIRUT*) indicated that if they were asked to draw all the details of the stellar parallax it would be much more helpful than connecting several dots and doing some adjustments on a computer.

Real-like experience: Four students from the *SPIRUT* group found it engaging finding distance not to a hypothetical star but to a real star, Sirius. They felt that they learned something that was not "fake" (PR03).

Interactive tutorials are perceived to be more helpful: students who worked with *LTs* expressed their wish to use a computer version of the tutorial instead. They argued that being able to interact would help them better with understanding of the concept. Only two out of 24 students who worked with *SPIRUT* preferred to work with a paper-based tutorial instead.

Bonus learning: another interesting aspect of *SPIRUT* was discovered. While students were thinking about one concept, they made unexpected connection with another concept of seasonal change of the stars. By observing how a planet was moving in its orbit, students commented that they noticed how stars are changing with the change of seasons and that must be why we see different constellations at different times of the year.

Animations can be deceiving or make students believe they understand the concept better than they do: students, who did not show a significant gain on the post-test, mentioned that they liked the animations because it made a concept clear. Without guessing whether they were doing things right or wrong, they could see the correct answer.

IV. Summary

During the interviews it was discovered that while most students learned something, very few learned a great deal. It was found that 35 out 36 students had a clear understanding that by observing an apparent shift of a star, the distance to it can be found. However, they were not so sure whether this apparent shift can be observed from different sides on the planets or it can be done from opposite sides of the Earth's orbit.

Many interviewed students who worked individually thought that they would perform better if they were to work with partners. At the same time, many students who worked with partners originally wanted to work individually. Additionally, 10 students identified themselves as independent learners. However, 15 out of 18 students who were randomly assigned to work with peers enjoyed the experience.

Students from the *LTs* group thought that they would learn more if they had immediate feedback to their answers. Twelve students mentioned Mastering Astronomy computer-based interactive exercises (currently used in the course) where they receive an immediate feedback whether the submitted answer is correct and in case if it is wrong they have an option of using a hint. Students from *SPIRUT* pointed out that feedback and hints were very helpful to them during the exercises. Students have come to expect a certain way of learning.

SPIRUT students found that starting exercises with a hard question was too difficult for them. They were not able to perform the required calculations to find a distance to a star. Ten out of 24 students would rather go through all parts of the tutorial from easy to difficult level in a step-by-step manner.

Students' most vivid memories of *SPIRUT* were interaction with the ships. Students readily and accurately described how by dragging a ship closer or farther away from the shore they observed its apparent shift against the background of the stars. Nineteen out of 24 students mentioned that they used this recollection while they were answering a question during the interview about apparent shifts of two stars.

Participants who performed better and worked either individually or with peers commented that the tutorials were hard, intimidating, and or confusing in the beginning but by the end they were able to see how all the pieces came together. They felt that they understood the concept and overall felt satisfaction after working with *SPIRUT*.

Students preferred multiple-choice to essay questions because they received immediate feedback that helped them with assessing their own knowledge. While essay questions required more time and deeper thinking they were not sure whether they answered a question correctly and as a consequence kept working on the exercises in the dark.

Thus, based on the analyzed students' responses in the interviews a certain pattern in students' learning behavior was revealed that suggested an answer as to why *Individuals* outperformed *Peers*.

Peers, students who were paired up with their peers, liked sharing ideas and learning from verbalizing their own thoughts. As a result of the frequent negotiation of answers, the students liked essay type questions that helped them to clarify their own thinking. They liked the ability to navigate in the computer-based environment skipping not-interesting, redundant, or hard parts.

Individuals, students who were forced to work independently, learned a great deal from the tutorials but found them hard; the students relied on feedback to assess their actions and were looking for hints to help them to find right answers. As a result of not being able to share ideas with partners, they preferred multiple-choice answers where there were fewer chances to lose points and a good chance of guessing the correct answer. *Individuals* had to put more effort into their own work and this may be a key reason for why they were able to perform better than *Peers*.

CHAPTER 5

DISCUSSION AND CONCLUSION

This Chapter presents a discussion of the findings, followed by recommendations for future studies, and recommendations for how to make the best use of the tutorial in a class setting. Before the details and interpretation of the findings will be discussed, the results for the research questions and hypotheses will be reviewed and additional findings will be briefly enumerated and highlighted.

I. Review of the Research Questions and Findings

The study was designed to answer the following research questions and show levels of confidence in the hypotheses listed below.

1.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

Hypothesis: Students who were involved in active learning through their direct interactions with the computer-based tutorial by means of a set of tools and were exposed to visual static and animated representations of the concept, would construct greater knowledge of the stellar parallax concept.

This hypothesis was supported by the results of the study. A group of students who worked with *SPIRUT* constructed greater knowledge of the concept.

In the study presented in this thesis, students had an opportunity to increase their knowledge of the stellar parallax concept by either working on computer-based tasks or by engaging in a paper-based assignment. Comparing the two types of treatment under these unsupervised or unguided conditions, without direct input of an instructor (Ardac & Sezen, 2002), it was found that the computer-based tutorial was more effective at increasing the understanding of the stellar parallax concept as a result of student's engagement with the tutorial.

Also, as the study showed, even though the two treatments were very different from each other, students in both groups benefitted from the treatments. In paper-based tutorials students benefitted mostly from working in a step-by-step manner, answering provided questions (from easy to hard), performing simple calculations, drawing lines on diagrams, and finding the shift of a star on multiple provided sketches. It is interesting that while students could not recall any details of drawing exercises, they remembered the overall step-by-step procedure that they found to be very helpful. One student recalled, "Probably doing the beginning steps of that packet [was most helpful]. Where it's more basic and then as you go through it they keep adding more to it. And just putting everything together step-by-step helped me learn it more than just being thrown into it like 'oh this is what you were supposed to do,' and this just breaks it down a little bit more" (PLT 04B). This comment is not surprising and fits well with previous studies (Vygotsky, 1978).

One of the features of *SPIRUT* is that it starts with a "real-world application" of a situated learning environment where students are astronomers who have a task in hand to determine a distance to a star based on simulated observations of stellar parallax. While importance of realistic setting for students' motivation was discussed by Gibbs (1988), Shank, Berman, and Macperson (1999), Marschall, Snyder, and Cooper (2000), to name but few, this opening scene created some confusion and shook students' confidence: "I remember that one [the first step in *SPIRUT*] and it was really hard. We started with that one and we got the very first question wrong. And it said that maybe you should try the first one before going on to this. I remember that after that remembering that this is going to be a rough lab" (PU 01B). While on the one hand it is possible that frightened students lost their confidence and performed worse than they would otherwise, it is also possible that the realistic and intimidating appearance of the application forced students to think more seriously about the work in hand and made them invest more effort in their learning. The way students process information depends on how they perceive this information (Solomon, 1983), "it is the effort demanding activities that produce better recall, comprehension, and inference-making" (Solomon et al., p. 44). Consequently, students in this study who persevered and invested mental effort in their work thought that it was very satisfying and motivating.

In addition, the use of a situated environment was noted by students as a learning stimulus, "I really liked how it wasn't just giving you a scenario where it listed a question and gave you data point. I liked how the simulation asked you to find the data and gave more real world application for parallax, you know I'm not going to go out and I can't really do this on my own, but it's kind of the idea that I could, if I really had the accurate tools and the instruments, I could figure out the distances to those distant objects and I thought that was really interesting and engaging. Although I found it hard to get the data accurately, I liked the idea behind that" (PR 13B).

Both tutorials had enough strong points that it enabled students to better grasp the concept, and both tutorials provided enough new exercises to solidify and build upon the knowledge constructed during the lecture part (prior to the pre-test and treatment).

1.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

Hypothesis: Students who were involved in active learning through their direct interactions with the computer-based tutorial by means of a set of tools and were exposed to visual static and animated representations of the mathematical representation of the concept, would perform simple mathematical computations and remember facts underlying the stellar parallax concept better than students who were performing similar calculations in the paper-based tutorial.

The results of the study did not support the hypothesis. Students from both groups significantly improved their mathematical skills on the post-test. Also, students who worked with the paper-based tutorials retained their skills slightly better after an eight-week period while students who worked with *SPIRUT* received similar scores to what they had on the pre-test.

Even though, the result was unexpected, it is possible that a vast amount of interactive exercise with emphasis on several aspects of the concept, made students disperse their attention. It is also possible that students' cognition was overloaded with multiple-representations of the concept that resulted in cognitive overload and precluded learners to perceive information in more efficient way (Mayer & Moreno, 2003).

At the same time, there were students who improved their mathematical skills as the result of their interactive experience with *SPIRUT*. One of the students commented, "…because you could see what happens to the calculations when you brought the ship closer or farther out. You could see if the numbers increased or decreased so it helped me as a person who doesn't understand math too well" (IU 23). By observing the actual change in the objects' positions on the screen and how it resulted in a change in values, the students were able to mentally connect the size of the shift with its numerical representation. A study conducted by Suh and Moyer-Packenham (2007) found similar results. Dual coding of verbal and non-verbal mental representations is most likely responsible for helping those students who at first struggled with the mathematical part of the concept.

It is also possible that whether students improve their mathematical skills and memorize basic facts or construct greater conceptual knowledge of the scientific construct depends not only on environment by on a combination of the environment and learners' prior knowledge, ability, and aptitude. This premise was suggested and investigated by Ardac and Sezen (2002).

These unexpected findings suggest that additional investigation is needed to find an answer to this new question.

1.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions with a computer-based or a paper-based tutorial?

Hypothesis: Students who were involved in active learning through their direct interactions with the computer-based tutorial by means of a set of tools and were exposed to visual static and animated representations of the concept, would learn how to reason their answers and transfer knowledge better than students who worked with the paper-based tutorial.

Results supported the hypothesis. The study revealed that students who worked with *SPIRUT* were able to apply their constructed knowledge to unfamiliar situations and find correct solutions.

Students who worked with *SPIRUT* had a significantly higher conceptual gain of the concept compared to the students who were given the paper-based treatment. These findings can be understood in terms of the dual coding theory described in Chapter 2. As a result of being able to see multiple representations of the concept: text, visual animated representations, interactivity, and scaffolding in a form of the tutorial adaptive features and feedback. Interviews indicated how each of these features of *SPIRUT* helped students to construct greater knowledge of the concept. In three interviews students mentioned that provided explanations divided into smaller sections and included on different screens helped them to process information better, “ I liked this because it kind of took us step by step through the whole stellar parallax idea. And then at the end it just pulled it all together, and really reinforced what you went through in the packet" (PR 04B).

Interactive exercises helped students connect constructed visual mental models with motor tasks, which resulted in deeper perception of and connection with the environment. This synthesis of logogens and imagens creates a more elaborated and involved mental model of the concept. In addition, a motor task together with a haptic perception (a feel of touch) of the world that is responsible for non-verbal coding (Pavio, 1986). Even though students did not "feel" objects (Sadoski & Paivion, 2001) directly in the sense that they did not walk on the sand on the shore, or that they physically pushed the ship farther into the sea, they felt as they did through manipulations of the tools as they observed a change taking place. This combination of visual and motor tasks, the dual coding process, also resulted in creating a stronger imprint on students' memories of the performed exercises, "The biggest recollection and the best example that I had was honestly in that lab… there was that exercise on where you are standing on a dock and you see a ship out at sea, and depending where you are standing on the dock and depending where you move across the dock, the ship is going to be in a different part of your vision" (PU 01B).

This would also explain why students in the *LTs* group did not have such vivid memories of their drawing exercises. They remembered only what they were already familiar with and this stuck in their memories. That would explain why when asked to sketch the stellar parallax model, several *LTs* students were sure that the parallax can be observed from two sides of Earth and not from two positions of the Earth in its orbit. In their visual memories they remembered a circle and they knew that they were observing the parallax from Earth, thus the circle they drew lines from in the lecture tutorial must be Earth. There were not enough visual and motor tasks to build a strong memory and since students held on to what they already knew about observing from Earth and their memory of a circle, it resulted in erroneous representation of the stellar parallax model. In addition, *LTs* lacked the immediate feedback that SPIRUT provided to students. When students clicked and dragged the ship, the tutorial responded with an immediate change on the screen, which resulted in coding this information into more imagens, but students in *LTs* group did not have such a visual response. They could not create as many imagens.

Therefore, despite some drawbacks that students referred to in their interviews regarding some specific features (or lack thereof) in *SPIRUT*, the fact that *SPIRUT* made effective use of multiple representations in conjunction with the students' ability to interact with them directly resulted in an overall increase in learning the stellar parallax concept.

2.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are aided by interactions in learner-controlled or program-controlled *SPIRUT*?

Hypothesis: Students who worked with program-controlled *SPIRUT* would construct greater knowledge of the stellar parallax concept as a result of *SPIRUT*'s tailored path for the students based on their actions.

The hypothesis was not supported by the results of the study. Students in both groups constructed comparable knowledge and significantly improved their conceptual understanding of the stellar parallax. Although, it should be mention, that students from both groups did not retain their knowledge.

2.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are aided by interactions in student-controlled or tutorial-controlled SPIRUT?

Hypothesis: Students who worked with program-controlled *SPIRUT* would learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better as a result of *SPIRUT* tailored path for the students based on their actions.

The hypothesis was not supported. Students from both treatment groups significantly improved their mathematical skills and were able to memorize basic constructs of the stellar parallax concept. Eight weeks later, however, this skill was lost again.

2.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are aided by interactions in student-controlled or tutorial-controlled environments of SPIRUT?

Hypothesis: Students who worked with program-controlled *SPIRUT* would learn how to reason their answers and transfer knowledge better as a result of *SPIRUT* tailored path for the students based on their actions.

This hypothesis was also not supported. Students in both groups improved their reasoning skills during their work in learner controlled and program-controlled *SPIRUT*.

It can be suggested that since there was no significant difference in gain between students who worked in the learner-controlled and the program-controlled environments and that since both groups almost equally improved their scores on the post-tests, that this type of scaffolding does not make a significant difference in students’ leaning achievements. While the distributions of the students’ scores were virtually identical for both environments, it is still possible that with a different design on the tutorial, the scaffolding would be more noticeable. However, it can be concluded that such a scaffolding effect would be of much lesser importance than the scaffolding effects that can be obtained from providing multiple representations. Nonetheless, our inability to discern between the two groups should be viewed as a limitation of study (related to the sample size) and perhaps improvements in the tutorial would remedy this situation.

Regarding potential improvements in the tutorial, students who were assigned to the learner-controlled environment mentioned that they felt lost while navigating the tutorial, “We had control over where we could go we didn’t really know what we didn’t really know. We got the answers wrong, it didn’t tell us to go back or forth we would just have to do it. And I think that if we would have the answer wrong, it would have taken us back and we would have to redo something, then it would have helped us more” (PU 14) This sense of loss is also visible in Fig. 4.40 where it can be seen that the students ranked their overall experience of their interaction with the tutorial higher in the program controlled version than in the learner’s controlled version.

The tutorial can be adapted to show students exactly where they are in the learning process, and what they are expected to learn from a certain section. If such a tutorial redesign would be effective, then this should show up as an improvement in the scores of the students assigned to the learner's controlled environment, thereby outperforming their peers assigned to the program controlled environment. Whether this would actually end up being the case will have to be gleaned from a future study.

3.1. Do students construct greater knowledge of the stellar parallax concept when their learning processes are supported by working with a partner or working independently?

Hypothesis: Students who worked with a partner would construct greater knowledge of the stellar parallax concept as a result of their cooperative work with the peer.

This hypothesis was not supported. The anticipated superiority of the peer instruction model over students working individually was not borne out by the present study. In contrast, it was found that working individually was actually beneficial in mastering the stellar parallax concept, independent of the type of treatment that was used to learn the concept. The difference between the two groups (individuals versus paired students) was noticeable, but not significant at the 95% confidence level.

3.2. Do students learn how to perform simple mathematical computations and remember facts underlying the stellar parallax concept better when their learning processes are supported by working with a partner or working independently?

Hypothesis: Students who worked with a partner would perform simple mathematical computations and remember facts underlying the stellar parallax concept better as a result of their cooperative work with the peer.

The hypothesis was not supported. Moreover, students who worked independently preformed slightly better than students who worked with peers.

3.3. Do students learn how to reason their answers and transfer knowledge better when their learning processes are supported by working with a partner or working independently?

Hypothesis: Students who worked with a partner would learn how to reason their answers and transfer knowledge better as a result of their cooperative work with the peer.

This hypothesis was not supported. Results of the investigation revealed that students who worked independently improved their reasoning skills and were able to transfer their constructed knowledge of the concept to unfamiliar situations better than students who worked with peers.

Perhaps the most surprising finding of the study was that paired students did not outperform their counterparts that were given the treatments on an individual basis. In fact, almost the opposite was found, namely that individuals outperform their paired peers with the caveat that this difference did not quite reach the $p = 0.05$ significance level. This finding appears to contradict the existing paradigm regarding peer instruction. According to the findings of a number of social scientists, it has been learned that peer work in pairs enhances students learning processes (Mazur, 1997) as a result of different types of student interactions and collaboration (Bransford, Brown, & Cocking, 2000; Duffy et al.1998; Palincsar, Brown, & Martin, 1987; Schwarz & Linchevski, 2007). Whether students take on the role of leader or participate in discussion of equals, they learn through verbalization of their thoughts (Mazur, 1997; Webb 1982) and through negotiation of meaning (King, 1998). However, in this study we found that students did not necessarily benefit from working with peers, which in retrospect may not be that surprising. There are a number of studies suggesting that peers do not always benefit from interaction – the nature of interactions can lead to positive as well as to negative results (Ge & Land, 2003; Pilkington & Parker-Jones, 1996). The conducted interviews revealed many clues as to why this finding might have materialized. The following describes these clues; in the part of this chapter where future studies are described it will be detailed how the hypotheses that can be based on these clues might be tested in new studies.

While there are a number of studies showing that learners do not necessary benefit from a particular type of interaction with peers (Hooper, Temiyakarn & Williams, 1991), there was no finding that indicated that students with high attitude and aptitude did not benefit from working with peers with the same achieving qualities. Some researchers predict positive outcomes from peer interaction in heterogeneous groups (Webb, 1982), others have an opposite opinion suggesting that homogenous groups are better suited for students learning from each other (Hooper 1992). In this study, the interviewed students who found themselves working with peers who had similar attitudes toward learning and prior knowledge did demonstrate significant gain in understanding of the stellar parallax concept following the treatment. “We were essentially able to have an open discussion with each other on how we felt on how the next move goes and whereas like I forgot a very important step he would have remembered it and so were like we were able to pick up the slack for each other. I thought working with a partner was awesome” (PR 01B). In interviews it was mentioned that explaining certain aspects of the concept to their partner aided them in developing a better understanding themselves, just as could be expected based upon the literature on peer instruction (Hooper, 1992; King, 1998; Solomon, 1984). “I think it was pretty good for me because first I learned by seeing it and then I learned by teaching it so- me teaching it to them was really beneficial to me as far as learning it” (PLT 06A).

Having said that, in pairs where one of the students had initially a better understanding of the concept and became a leader, another student would (or could) follow without gaining additional understanding of the concept s/he already had before the exercises. As mentioned above, in this situation leaders would find working with a partner beneficial to them as a result of becoming an instructor to their partners; through their own teaching they would achieve a better understanding of the concept which would show on their scores (King, 1998). At the same time however, their partners either would not improve their knowledge resulting in a lower or non-significantly improved score on the post-test, and/or they would receive significantly lower scores on the retention test.

The study demonstrated negative outcomes of a random pairing up of peers. Here is an example of how one of the students from a homogeneous pair felt about his "collaboration" with his partner, "I have no idea if he was doing it right. He would look at it and be, like, 'well this might be right,' and he would, do something, but…he wasn't super… we couldn't really, help each other with definite answers cause neither of us really knew definite answers. But I think I probably was more helpful about it than he was" (PU 14A). This demonstrates that learners with lower prior knowledge will not achieve better learning outcomes from working with peers.

Interestingly, students who worked in pairs and performed worse than their partners on the post-test commented that they really enjoyed working with their partners and that they both contributed equally and felt that both of them had the same level of understanding of the concept. However, from their answers to follow-up questions, it was clear that these students required more time to think about the problem and wanted to go back to see their previous answers on several occasions. The same students wanted to spend more time to think about the questions and they commented multiple times that their partners explained the answers to them on many occasions. One of these students shared, "Sometimes, if we just have to move something around, maybe the partner doesn't prefer to do that. Maybe I am confused about one question but he gets it and he wants to skip over these questions to move to the next one. But I am still confused on this question, so maybe I need more time. But he may think it is kinda (sic) wasting time, his time" (PU 04).

These findings also bring up an issue of perceived self efficacy discussed by Bandura (1982), "When beset with difficulties people who entertain serious self doubts about their capabilities slacken their efforts or give up altogether, whereas those who have a strong sense of efficacy exert greater effort to master the challenge" (p. 123). Self-efficacy as an important factor in determining by a student how much of mental effort he or she needs to invest in learning. The scientist suggested that learners with high level of perceived self efficacy may over estimate their abilities and thus not invest enough mental effort which can hinder their learning. Evidently, the results of our study support this statement.

The above paragraphs provide some insight into why some pair groups might not have performed as well as could have been hoped, but they do not provide us with any insight as to why individuals did better. The above passages only illuminate why some pairs might have done just as bad as some individuals. Clearly, it would be interesting to learn the exact reasons why individuals actually appear to have done better.

One potential reason for individuals doing better might be that some people like learning on their own, while others learn best by bouncing ideas off of each other. In other words, does learning preference correlate with outcome? Some of the interviews provided insight into the learning dynamics of the paired individuals, and perhaps these observations, discussed in the following paragraphs, hold some clues as to why pairs fell behind their individual counterparts. These findings should be viewed as clues, the sample size in this case is clearly too small to draw any conclusions with certainty.

For instance, students who liked working with a partner shared that they were learning the concept better while they were explaining it to their peers. It was mentioned by 10 out of 18 students in the *SPIRUT* groups and by two out of six students in *LTs*. This would be a point in favor of correlating learning preference with eventual outcome.

Similarly, students who did not like working with peers pointed out that the reason for it was that they felt being hurried along by their partners and that they did not want to hold back their partners. In this case, the pair dynamics adversely affected the learning outcome.

In addition, students from two *LTs* pair group and two students from *SPIRUT* felt like they did all the work and basically worked alone. In such cases, learning preferences might have stood in the way of optimal learning experience as it would appear that students who felt hurried along by their partner would have benefited from either working individually, or by being matched to a student of more similar capabilities, or by being matched with a student who was perhaps not in a hurry. Such mismatches are more unlikely if students get to choose and pick their partner for themselves.

Based upon these answers, it would appear that the peer instruction model does not always work. The pillars of peer instruction are that even in the case of unmatched individuals, the better student would gain knowledge from explaining the concept, while the other student would gain from listening. The latter statement does not appear to stand up to scrutiny in this case: either the mismatch was too great, or the quality of explanations not good enough to be truly helpful. It should be noted that in standard peer instruction, the instructor would step in at this point, something which could not be done while doing these tutorials.

The above statements rely quite strongly on the students' input based upon the students' perception. However, these perceptions do appear to correlate with what was learned by simply looking at how well the students did on the post-test compared to the pre-test. These numbers were shown in Tables 4.40 and 4.41 (Chapter 4, page 202).

Table 4.40 shows that students who described themselves as leaders did quite well (2 ‘A’ gains’ and 1 ‘B’ gain), whereas students who described their participation in the tutorial as ‘bouncing off ideas’ were passive followers. In fact, the group that did the worst consisted almost entirely (5 out of 6 students) of students who liked to bounce off ideas. Their test results show that for these 5 students, bouncing off of ideas actually involved doing very little, but rather let their partner take care of business, “I mean she definitely pulled her weight. But, I feel, like, cause some of her questions I had to answer. I mean they were probably wrong, but I still feel like it was, sixty and forty percent” (PR 14A). Such students would most likely benefit much more from being assigned to do their work on an individual basis, and it probably would also have helped their partner who worked as an individual for most of the time anyways. Therefore, the results shown in this table suggest that students who like working in pairs because they are focused on learning benefit from the paired assignment, whereas students who like working in pairs because they have the opportunity to exchange ideas might not benefit unless they are willing to put in the effort required for learning. If they do not put in such an effort, then they actually gain negative benefit from being paired up.

Table 4.41 shows that individuals who liked working alone did very well on the assessment test, with 5 out of 6 students receiving the highest gains (‘A’ group). As one of the students pointed out, “Because I just work by myself most of the time. I feel confident in myself I guess” (IR 25).Thus, these tables are consistent with the notion that on average, for all students that partook in the tutorials, individual grouping would be better than assignment in random pairs.

There appear to be many factors that can adversely influence the learning dynamics of a paired group. The difference in aptitude has been described, but other factors were identified as well during the interview stage. For instance, one of the students could want to speed along through the tutorial because of prior engagements such as a favorite television show starting. The other student then would feel hurried along, and not spend as much time on certain exercises as s/he would have liked to. As one student mentioned, "It depends on the person: if they actually care or not about the tutorials" (ILT 06). One student might be temporarily distracted, one student might be having an off-day, and one student might simply view having a partner as a wonderful opportunity to not work very hard without affecting his or her grade. Such differences in attitude are likely to affect the outcome for the pair adversely compared to the outcome for individuals. After all, individuals will not be hurried along, nor can they hide behind a partner when their attitude is such that learning is secondary to getting a good grade. Here is a part of the recorded interview that demonstrates a student's reflection on her interaction with the computer-based tutorial:

Student: "[…] Just a lot of them moved around, the motions I guess. I know they were trying to be helpful and explain things but I was more worried about my grade than about understanding it."

Interviewer: "If no grade would be involved do you think you would want to learn more?"

Student: "Yes"

Interviewer: "If no grade was involved, why would you want to learn about it?"

Student: "Because it's what we look at in the sky every night." (IU34)

A concrete example is the case for pair groups 5 and 8. While the fact that both these students worked with partners and lost points on the post-test can be purely coincidental, it is possible that these students did not invested enough of mental effort (Salomon, 1983) in their tutorial and thus did not contribute in their learning process as much as their partners did. This would directly tie into the dual coding mechanism: if students do not put in much mental effort, then they do not create new logogens and imagens (Paivio, 1986). Therefore, it can be suggested that if students are not actively involved in the exercises (observe and interact with a computer or perform a paper-based tasks) then they do not put enough mental effort and thus they cannot process (or code) the information. Without paying close attention to what is taking place on a screen, students do not receive as many visual representations, and they do not participate in coding and analyzing. If both these particular cases are compared to each other, it can be also suggested that a student from the *LTs* group had less visual and motor tasks to perform in the first place compared to the computer-based tutorial students. Either way, whether mental effort necessary for creating imagens and logogens is lacking by the students' choice, or by the design of the tutorial, it certainly would be interesting to investigate (quantitatively) the potential correlation between learning outcomes and the opportunity to create imagens and logogens.

Again, it would appear that all these clues and indications lead to the conclusion that students should not be assigned to work with partners randomly. A selection of partners should be based on several criteria: level of interest in learning the material, same goal of achieving a higher grade, familiarity of partners, and their knowledge of the concept, “it would really just depend on the other person’s knowledge. If he didn’t know anything I would rather just work by myself” (IU 36).

While in a classroom with the presence of an instructor the latter is not an insoluble problem, in an environment not controlled by an instructor two peers that do not naturally benefit from each other’s knowledge are more likely to not make much progress. Therefore, one of the further directions of the research should be to investigate whether the peer instruction model works in unsupervised environment, and whether students who choose their partners based on familiarity with each other can reap the benefits from collaborative work on assignments. Another direction for future research is to investigate how students should be paired up to benefit the most from their collaboration.

In summary, there are many potential reasons why individuals outperformed their paired peers, but at this point there is not a single reason that can be identified as the major reason. All indications are that unless the paring of the students is such that they are similar in both aptitude and attitude, that the pairing might not carry additional benefits with it other than for the student with the higher aptitude and better attitude. As mentioned, these are clues gleaned from the interviewing process and more studies will have to be conducted as to the pairing process relevant to unsupervised tutorials.

II. Other Findings

As our findings showed, students who worked with the computer-based and the paper-based tutorials significantly improved their knowledge of the stellar parallax concept as a result of the treatments. Students in both groups received significantly higher scores on the post-test. This confirms previous theories and empirical findings starting with works of John Dewey (1938) and Lev Vygotsky (1978) that students learn better when they are actively involved in learning processes through a number of activities and when these activities are supported or mediated either by guidance of a more capable person, a tutor (Heift & Nicholson, 2002; King, 1998; Mitrovic, 2003), collaboration with peers (Vygotsky, et al; Webb, 1982; Mazur, 1997; Duffy et al, 1998), or scaffolded by tools in a computer-supported environment (Hmelo-Silver, Duncan, & Chinn, 2006; Davis & Linn, 2006; Kolodner, Owesby, & Guzdial, 2002; Azevedo, Cromley, & Seibert, 2004). Of course, part of the improvement is likely to be ascribable to the fact that students go over the same concept one more time. However, this study was not setup to measure such a repetition effect and therefore it is included in the discussion.

However, irrespective of the treatment: pairing, and administration of the tutorial, retention of the learners' knowledge of the stellar parallax concept dropped back to levels similar to those measured directly (on the pre-test) after the brief 20-minute lecture. This does not imply that the same aspects were retained or forgotten as compared to the pre-test, it merely implies that the overall understanding dropped to pre-treatment levels. This does not equate to a failure of the treatments as it is possible that students would have dropped to well below the pre-treatment levels without the treatment, but it does show that a considerable amount of knowledge was gained and then not retained.

Neither treatment was successful at maintaining post-test gains when re-measured eight weeks later.

First, even though both groups showed similar results on the retention test, this may actually refer to a limitation of the study. The three assessments tests covered the same questions, albeit in different orders. It is possible that when students answered those questions they were recalling questions and their answers from previous assessment and did not think about the actual concept. For example, during interviews when students were asked to explain their answers to some of the questions they shared similar reaction: “Oh this one. I remember it. I do not remember what my answer was. That was a tough one” (ILT 04).

Second, what is clear is that a considerable amount of knowledge is lost during an eight-week time span when the learned concept is not being revisited, as was the case in the present study (by design). This does not normally correspond to the actual classroom settings where students are being quizzed on the learned concepts after a shorter time span. Of course, preparing for such a quiz would again increase the overall level of knowledge and understanding of the learned concept. The present study was not designed with this in mind, but it could easily be adapted to actually measure retention under such more normal classroom conditions. This way, the influence of the treatment on retention might be brought to the foreground.

Therefore, while the levels of retention were low, and while there was no correlation between those levels and the treatment given, and the conditions under which the treatments were being administered, it is too early to draw the conclusion that treatments will not be successful at improving long-term retention.

The conducted interviews revealed that immediate feedback was important to students. This finding applied to students subjected to either type of treatment. Those whose interactions were scaffolded by feedback commented positively on it, those that did not receive feedback commented on its absence. Students linked the availability of feedback directly to their ability to learn and understand the concept.

Feedback was a very important design feature of *SPIRUT*. For instance, when students picked a wrong answer a line of text or animation would be shown to hint students toward the right direction to allow them to correct themselves. Students repeatedly pointed out that it helped them a lot in learning the concept when going through the tutorial. Perhaps, it was the most useful form of scaffolding implemented in *SPIRUT*, “It would tell you if you got it wrong and it might show you an animation or something, and I thought that was really helpful” (IR 02).

While many students commented that they liked essay questions because they, ”made me think more” (IR 02), 11 students pointed out that they would rather do multiple-choice questions because they knew right away whether they were right or wrong. “The multiple choice questions were helpful because you could make sure you were doing the right thing. I didn’t really like the ones where you had to type it out because it wouldn’t tell you if you were right or wrong. So I could be completely wrong on those and I wouldn’t know” (IU36).

Students who worked in *LTs* group suggested that lack of feedback was a main disadvantage of the paper-based tutorial, “you could do all exercises and would not know that you got them all wrong” (ILT 04).

Another form of scaffolding used in *SPIRUT* was its adaptive response feature, its ability to react on students’ input: mistakes and repeated mistakes. When students answered a question incorrectly they would be provided with a line of feedback hinting them into the right direction, a second mistake would generate a relevant animation or interactive exercise and if these failed to help, the tutorial would take users to a part of the tutorial that would help students to find out why their selected answers were incorrect (in the program-controlled version), or the tutorial would tell the student s that they might want to (re)visit a certain section (in the learner’s controlled version).

In conclusion, students were very appreciative of receiving feedback. The students that were interviewed indicated that feedback helped them in their learning. Note that the study did not actually measure whether feedback did increase their learning; at this point we can only conclude that students perceive that feedback does indeed help them.

III. Discovered Themes

In this part some noteworthy observations will be described and discussed. These observations do not amount to major findings, but they are striking enough to be listed.

One of the unexpected discoveries concerns the students' view of the place of computers in the learning process. Even today, some students who think of themselves as proficient computer users do not see a computer as an equal or a superior to a paper-based exercises learning tool. Some students perceive use of computers as more of a demonstration of a phenomenon that can equip students with visual representations, help to envision a phenomenon from different points of view (the Moon from Australia, for instance) but they do not see it as an equal learning tool to paper-based exercises. "The computer was nice to see the difference, like you could move a tool, and be like, ok if it is 12 o'clock here, like in other words. If you changed it, you physically saw the change and then you could move it back. But with paper more so, you have to think about it, there is more thinking to completely understand it and before you can do it" (PU06B). It is interesting that this student was working with a partner in *SPIRUT* environment and his measured gain dropped on the post-test. During the interview, the student was not able to give a clear explanation of the stellar parallax (he did recall that it deals with distances and shift), sketch the model or write a formula down.

Another student, who worked in *LTs* thought that he would not be able to learn the concepts when using a computer just as well as he would learn it otherwise, "I just like being able to do it by hand than by just clicking, and choosing, or even drawing different things. I mean the questions would be the same so that wouldn't make the difference on the questions part. But more on the diagrams- especially like these sections where we're identifying the stars, you'd just be clicking. Whereas here I look at it more and determine it that way. I don't know, it seems different on a computer" (PLT 15B). It is interesting that this student uses a computer every day and he remembered the overall idea of stellar parallax, but he could not sketch the model correctly or recall the formula; he also thought that the lecture helped him the most in learning the concept. He could not recall any particular drawing exercises or visual representations from his work with lecture-tutorials until he saw the booklet again.

Also, the third interview might hold a clue to what these students really lack when using a computer as a learning tool. It is possible that they perceive information received from a computer as not being tangible, something that they cannot hold on to and lock it in their heads. In addition, some students wanted to have some tangible evidence of their work, such as their printed out answers, report of their work, formulas, etc. “I wish we had something to take… like a paper to take back from it. I usually like to have papers and to keep notes cause I think if we have those, a concrete paper, then it’s like…okay, you can work through this paper, but then you can also work through it on the computer, that way you can take something with you and that way you have all those formulas still and you still have all of the things that you learned from the simulation in a hard copy” (PR 01A).

IV. Limitations of the Study and Future Development

The directions for future studies are based both on the findings of the present study, as well as on the (inherent) limitations of this study. The limitations of the present study are summarized in the following as well as a brief discussion on whether the study should have been constructed differently from the outset.

IV.I. Limitations of the Study

- The overall sample size was restricted to the number of students taking the introductory class and that were able to complete all three assessment tests. This is not a limitation that can be overcome unless the study is conducted at a University where larger class sizes are the norm. The sample size directly relates to whether observed effects can be observed with high accuracy ($p< 0.05$). However, while it is an obvious limitation of the study, it cannot be changed by a different study design. For instance, increasing the sample size by spreading the study out over multiple semesters would invariably introduce comparison problems between the multiple populations.
- The computer-based tutorial was very good, but not perfect. It is only after having a multitude of students go through the tutorial that some limitations come to the fore. An example of this would be that students assigned to the learner-controlled environment reported feeling a sense of being lost while navigating the tutorial. This might have resulted in skewing the difference in performance between learner-controlled and computer-controlled environments.
- The study relied on assigning numerical values to the students' responses to essay questions. Even though a co-rater was used to minimize the effects of human error during the assessment and inter-rater agreement was obtained (answers are not just absolutely right or wrong but rather assessing them involves a lot negotiation and assumptions about what students really meant in their answers), rating essay questions will always result in a limitation of the study. Coding interview responses leads to similar limitations as the wording of the student might not necessarily correspond to what the student is trying to say, or what s/he is capable of.

- Some responses to interview questions might have been skewed as some students have a natural tendency of trying to please the interviewer.

- The timings of some of the treatments or assessment tests were dictated by room and students availability; this could have resulted in some students hurrying through some parts of the study so as to not miss their next appointments.

- The assessment tests consisted of the same questions, given three times. Such a repetition can result in students answering from memory, or students not giving it ‘their best’ as they have already given the correct answer in the past and do not feel inclined to put in the mental effort to do so again.

- One question on the assessment test was not exactly right for both types of treatment, and had to be thrown out for some parts of the study where the two treatments were compared. This reduced the power of the study compared to the intended study design.

- The study was designed to test and investigate certain issues, whereas some students only saw the tutorials as a recipe to a good grade, and were only taking it with the attitude of spending the minimal amount of mental effort on it while hoping for a maximum outcome. This might have skewed the premise that all students were more or less equal at the outset of the study; the study did not differentiate between students that were taking the tests and tutorials with the aim of learning, and those that took at in order to simply get it over with.

IV.II. How Could the Study Have Been Improved?

- Within the limitations listed above, the study was constructed adequately, but some of the study's elements could have been improved. As mentioned under the study's limitations, one question did not equally assess the computer-based and paper-based tutorial groups. Clearly, this question should have been different.
- The computer-based tutorial was not perfect. With fairly minor modifications, the students could have been given a better sense of where they were in the learning process, and what they were supposed to be learning. It would also have been helpful to have the tutorial commence with a brief clip of what learning tools were available to the students within the tutorial. The tutorial also encompassed many animations, but these animations could not be stopped or paused by the students. This frustrated the students, and probably interfered with their learning. In addition, based on the results of the study in conjunction with the dual learning concept, some exercises can be added to the computer-based tutorial to aid the students in constructing their knowledge through multiple representations at points where the present version of the tutorial only relies on single representations.
- Given the knowledge that has been gained from administering the assessment tests, some of the questions can now be reworded to result in a higher statistical reliability (Cronbach's alpha statistical test). The questions as given on the assessment test in the present study were selected and reviewed by a group of instructors for clarity and relevance, but not for statistical reliability.

IV.III. Future Investigations

All the findings, and limitations, of the present study lead to the following directions for future studies.

IV.III.I. Interaction of students with a computer-based tutorial.

Tutorials are designed with a particular learning outcome in mind, and tutorials are taken by real students that all have different attitudes and aptitudes. It would be interesting to see if all these differences nonetheless translate into some common interaction between the students and the tutorial during the learning process, and if so, how tutorials can be (better) designed to take advantage of such findings. In particular, one would like to know what exercises engage the students, and which ones help them learn. With this in mind, when *SPIRUT* was administered to 199 students, the following input was recorded. A video and audio recording was made of all students taking the tutorial in conjunction with their actual input on the screen and keyboard. In addition, an electronic weblog was kept of all students' answers to all exercises, as well as of their navigation through the tutorial. The collected material will be analyzed in the foreseeable future.

IV.III.II. Learner-controlled versus program-controlled tutorials.

As discussed in the preceding, our study could not differentiate between the learner-controlled and computer-controlled administration of the computer-based tutorial. After the modifications to the *SPIRUT* tutorial (as indicated previously) have been completed, this issue can be revisited. Sample size can be increased for this study within the limitations of the size of an intro class by assigning all students between those two groups (as individuals), without having a paper-based control group.

IV.III.III. Step-by-step approach versus starting in the middle.

The starting point of the computer-based tutorials was chosen to coincide with the amount of knowledge students could have learned from the in-class lecture had they understood most of it. This point was chosen so as to accommodate the very best students by not making them do the equivalent of busy work, while not standing in the way of the learning of the students who were not at that stage yet. Students who were in the computer-controlled navigation environment would be quickly redirected to parts corresponding to their level of knowledge; students who were in the learner’s controlled part of the navigation had to find their way to these parts through the (strong) hints that the program would give them.

While there are many considerations to take into account when allocating a starting point, and while some students commented on the choice of starting point, it is not clear whether such a choice affects the eventual outcome of the learning process. A study can be designed to investigate this aspect to see if an improvement can be found in the way tutorials are being administered as the common way of administering tutorials is to start at the (easy) beginning, and force students to make their way to the end. This is the premise behind the Mastering Astronomy tutorials. The proposed study would measure whether such an approach perhaps results in students losing interest during the easy stages, or simply running out of time, and as a result, not achieving the desired learning outcomes upon reaching the harder parts of the tutorial.

IV.III.IV. How does the choice of a partner affect the learning outcome?

The present study revealed that learning outcomes differed not only between individuals and groups, but also within the two students forming a peer group. In particular, it was found that some students who teamed up with someone of better aptitude or better attitude did not achieve very good gain in understanding. It is unclear whether this was due to the students' own capabilities, or to the fact that they were mismatched in the peer settings. A study that assigns pairings based upon students' performances on already administered quizzes and exams should be able to investigate whether pairing can actually negatively affect the learning outcome when a paring with a different individual could have resulted in a positive outcome.

A related study would be to investigate whether there would be a difference in learning outcome when, on the one hand, students pick their own partners and, on the other hand, partners are assigned at random. Both studies would potentially have important implications for the peer instruction model.

IV.III.V. Is the peer instruction model (still) valid in unsupervised settings for groups of two students?

The standard peer instruction model is not only based upon the students' interaction with each other, but it also requires an instructor who is present in the room to ‘break the gridlock,’ or to direct the peer group when it is headed in the wrong direction. Unsupervised tutorials in which students strictly work in pairs do not fall into this category, and the present study shows no beneficial effects from the pairings.

Therefore, one of the future directions for research is to investigate whether the peer instruction model can work in an unsupervised environment, whether students who chose their partners based on familiarity with each other can really benefit from collaborative work on assignments. Another direction for the research is to investigate how students should be paired up to benefit the most from their collaboration. A study can be designed to investigate whether pairing should be done under unsupervised conditions. The study design would encompass a group of paired students who complete a tutorial without the presence of an instructor, and this group would be compared to the learning outcomes of paired students who completed the identical tutorial in the presence of an instructor and in the presence of other paired groups completing the same task.

IV.III.VI. Do tutorials (computer-based and paper-based) aid in retention of the learned materials?

The present study was not designed to measure retention of the learned materials under normal classroom settings that involve students revisiting the materials during the process of preparing for quizzes or exams. With minor changes, the present study design can be used to actually measure the retention potential of the tutorials and normal classroom settings. A virtually identical procedure would be used, but a few weeks after the administration of the treatment the students would be subjected to a test that counted towards their grade, followed by a retention assessment test some 6 weeks later.

Hopefully, such a study would not only reveal a less precipitous drop in retention level, it might also be able to differentiate between the various treatment options.

IV.III.VII. Incidental learning: can learning outcomes be improved for students not interested in the topic?

Quite a few students complete a tutorial with their grade as the ultimate and immediate goal. Typically such students are not necessarily interested in the material, but rather they are simply completing a grade requirement in a required course. It is very possible that such an attitude adversely affects the learning outcomes as learning becomes incidental during such an approach, secondary to completing the task at hand.

A study can be designed where the participants are chosen from a group of students who are, by their own admission, not interested in the topic that they are learning. The study can then assess whether animations, video clips, and interactions with a tutorial would engage students to such a degree that learning becomes the first priority once again. This is a complicated study to design; in here we only mention the idea for the investigation, not the details for designing such a study.

IV.III.VIII. Feedback: real or perceived learning improvements in computer based tutorials?

During the interviews, students commented frequently on the feedback available to them, or unavailable to them. Yet, while the differences in learning outcomes between the computer-based tutorial that included feedback, and the paper-based tutorial that did not allow for a feedback mechanism revealed significant difference in learning outcomes, it was not clear how much of those differences could be ascribed to the feedback provided. In other words, while students do have the perception that feedback is beneficial to them, it might actually not be the case. One can envision a situation where a student gains a deeper understanding because the lack of feedback forced the student to come up with the solution by him or herself, without the aid of a hint. As with the previous suggestion for future study, such a design is beyond the scope of this thesis.

V. Recommendations

The following are a list of general recommendations for actually increasing the students' knowledge of the stellar parallax concept when it is being taught not as part of a study, but in normal classroom settings. These recommendations are based on the findings of the present study.

- The SPIRUT tutorial that was developed for this study turned out to be a very useful instruction tool that could be used by individual learners. As such, SPIRUT would be good to use in online classes. The computerized feedback in combination with the lack of time constraints should lead to a meaningful learning experience, even in the absence of an instructor.
- When designing classes, it is important to focus on retention. Students are entirely capable of forgetting considerable amounts of what they learned 8 weeks prior, and our study showed that this is where many gains can be made.

This retention issue should be an important consideration in deciding upon how many exams should be given.

- The natural recommendation that follows from this is that one should ensure that online classes are accompanied by short video clips where the concept is explained. One should not solely rely on textbooks or tutorials to do an equally good job.
- Both paper-based and computer-based tutorials require a large time investment (to create); especially the computer-based tutorials require a considerable investment in time and money. The recommendation is to only create computer-based tutorials if they will also be used for online instruction, and that they should be designed with online instruction in mind.
- Should a tutorial be used in instruction in a physical classroom, then in view of the different preferences of the students and in view of how students learn best, students should be given a choice of how to use such a tutorial. Students should be allowed to choose whether they want to navigate the tutorial on their own or with a partner, and if they choose the partner option, they should also be able to choose their own partner. Students should be given the option of whether they want to navigate through the tutorial themselves, or whether they want to be taken through it on a step-by-step basis.
- When administering a computer-based tutorial, students should be able to take home a printout or some other record of their learning. While *SPIRUT* actually created an electronic logbook of their entire tutorial session, this logbook was not given to students as its original intent was to use it for

interpretation of the students' learning. However, such a logbook would be helpful to students' retention and preparation for exams, as mentioned by students during interviews.

VI. Bloopers

The final part of this thesis consists of direct quotes taken from the transcribed interviews of the students. This section is merely to illustrate that no study can ever be designed perfectly, and that no teacher can ever reach all the students in a class or teach at the appropriate level for all hundreds of students. Enjoy.

- "I know if you put your finger in the air, and if you close one eye. I know that part because I thought it was cool."
- "Can you mathematically represent the relationship between the distance to a star and its parallax?" (the actual complex formula is: $d = 1/p$)
- "Okay well this is one AU, and then [pause]…that's as far as I can go"
- "There are parsecs and a lot of stuff"
- "Isn't it…something divided by something?"
- "D= p/a^3"

- “It had something to do with arcseconds but I was never able to get the math part of it down well.”

- “It’s like this really long equation divided by something. It’s kinda (sic) long, but I am not good at math. It is divided by 10 on the last part over here. Yeah, it’s like 4.52 times 10 or something like that. I don’t know.”

- “If a star was moving directly away from us, it wouldn’t shift…as much. So it could be going, I mean, I’m talking about a star, if a star could go really, really fast, but it was going directly away from us, we wouldn’t observe as much of a shift and if another star is moving very slowly but it’s moving, umm, side to side of us, we would see a lot more of a shift and we would assume that it’s going faster, so I mean it really depends on what direction it’s going in, what we observe and how far away it is.”

- “And then answering questions made me think through things rather than just kind of think through things. The questions made me be like- this is why, and have to explain myself.”

APPENDIX A

Why did the ancient Greeks reject the real explanation for planetary motion?

If the apparent retrograde motion of the planets is so readily explained by recognizing that Earth orbits the Sun, why wasn't this idea accepted in ancient times? In fact, the idea that Earth goes around the Sun was suggested as early as 260 B.C. by the Greek astronomer Aristarchus. No one knows why Aristarchus proposed a Sun-centered solar system, but the fact that it explains planetary motion so naturally probably played a role (see Special Topic, p. 50). Nevertheless, Aristarchus's contemporaries rejected his idea, and the Sun-centered solar system did not gain wide acceptance until almost 2000 years later.

Although there were many reasons why the Greeks were reluctant to abandon the idea of an Earth-centered universe, one of the most important was their inability to detect something called **stellar parallax**. Extend your arm and hold up one finger. If you keep your finger still and alternately close your left eye and right eye, your finger will appear to jump back and forth against the background. This apparent shifting, called *parallax*, occurs because your two eyes view your finger from opposite sides of your nose. If you move your finger closer to your face, the parallax increases. If you look at a distant tree or flagpole instead of your finger, you may not notice any parallax at all. This little experiment shows that parallax depends on distance, with nearer objects exhibiting greater parallax than more distant objects.

If you now imagine that your two eyes represent Earth at opposite sides of its orbit around the Sun, and that your finger represents a relatively nearby star, you have the idea of stellar parallax. That is, because we view the stars from different places in our orbit at different times of year, nearby stars should *appear* to shift back and forth against the background of more distant stars (Figure 2.34).

Because the Greeks believed that all stars lie on the same celestial sphere, they expected to see stellar parallax in a slightly different way. If Earth orbited the Sun, they reasoned, at different times of year we would be closer to different parts of the celestial sphere and would notice changes in the angular separations of stars. However, no matter how hard they searched, they could find no sign of stellar parallax. They concluded that one of the following must be true:

1. Earth orbits the Sun, but the stars are so far away that stellar parallax is undetectable to the naked eye.
2. There is no stellar parallax because Earth remains stationary at the center of the universe.

Aside from a few notable exceptions such as Aristarchus, the Greeks rejected the correct answer (the first one) because they could not imagine that the stars could be *that* far away. Today, we can detect stellar parallax with the aid of telescopes, providing direct proof that Earth really does orbit the Sun. Careful measurements of stellar parallax also provide the most reliable means of measuring distances to nearby stars [Section 15.1].

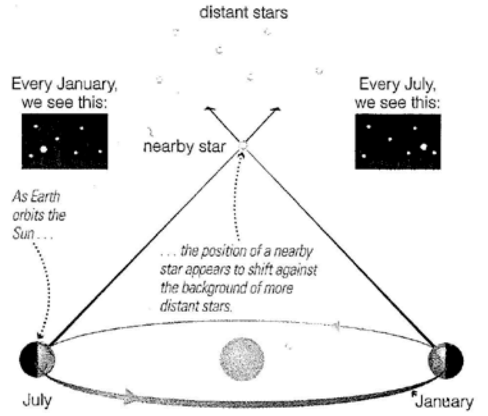


FIGURE 2.34 Stellar parallax is an apparent shift in the position of a nearby star as we look at it from different places in Earth's orbit. This figure is greatly exaggerated; in reality, the amount of shift is far too small to detect with the naked eye.

THINK ABOUT IT

How far apart are opposite sides of Earth's orbit? How far away are the nearest stars? Using the 1-to-10-billion scale from Chapter 1, describe the challenge of detecting stellar parallax.

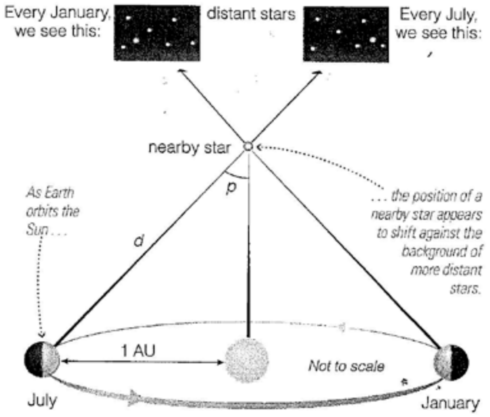


FIGURE 15.3 Parallax makes the apparent position of a nearby star shift back and forth with respect to distant stars over the course of each year. The angle *p*, called the *parallax angle*, represents half the total parallax shift each year. If we measure *p* in arcseconds, the distance *d* to the star in parsecs is 1/*p*. The angle in this figure is greatly exaggerated: All stars have parallax angles of less than 1 arcsecond.

Measuring Distance Through Stellar Parallax The most direct way to measure a star's distance is with *stellar parallax*, the small annual shifts in a star's apparent position caused by Earth's motion around the Sun [Section 2.4]. Recall that you can observe parallax of your finger by holding it at arm's length and looking at it with first one eye closed and then the other. Astronomers measure stellar parallax by comparing observations of a nearby star made 6 months apart (Figure 15.3). The nearby star appears to shift against the background of more distant stars because we are observing it from two opposite points of Earth's orbit.

We can calculate a star's distance if we know the precise amount of the star's annual shift due to parallax. This means measuring the angle *p* in Figure 15.3, which we call the star's *parallax angle* and is equal to *half* the star's annual back-and-forth shift. Notice that this angle would be smaller if the star were farther away, so we conclude that more distant stars have smaller parallax angles.

All stars are so far away that they have very small parallax angles, which explains why the ancient Greeks were never able to measure parallax with their naked eyes. Even the nearest stars have parallax angles smaller than 1 arcsecond—well below the approximately 1 arcminute angular resolution of the naked eye [Section 6.2]. For increasingly distant stars, the parallax angles quickly become too small to measure even with our highest-resolution telescopes. Current technology allows us to measure parallax accurately only for stars within a few hundred light-years—not much farther than what we call our *local solar neighborhood* in the vast, 100,000-light-year-diameter Milky Way Galaxy.

By definition, the distance to an object with a parallax angle of 1 arcsecond is 1 **parsec (pc)**. (The word *parsec* comes from combining the words *parallax* and *arcsecond*.) Because all stars have parallax angles smaller than one arcsecond, they are all farther than 1 parsec away. If we use units of arcseconds for the parallax angle, *p*, a simple formula allows us to calculate distances in parsecs:

$$d\,(\text{in parsecs}) = \frac{1}{p\,(\text{in arcseconds})}$$

For example, the distance to a star with a parallax angle of $\frac{1}{2}$ arcsecond is 2 parsecs, the distance to a star with a parallax

---

*Astronomers sometimes refer to the total luminosity as the *bolometric* luminosity.

angle of $\frac{1}{10}$ arcsecond is 10 parsecs, and the distance to a star with a parallax angle of $\frac{1}{100}$ arcsecond is 100 parsecs.

Astronomers often state distances in parsecs, kiloparsecs (1000 parsecs), or megaparsecs (1 million parsecs). However, with a bit of geometry, it's possible to show that 1 parsec is equivalent to 3.26 light-years (see Mathematical Insight 15.2). We can therefore modify the above formula slightly to give distances in light-years:

$$d\ (\text{in light-years}) = 3.26 \times \frac{1}{p\ (\text{in arcseconds})}$$

In this book, we'll generally state distances in light-years rather than parsecs.

Parallax was the first reliable technique astronomers developed for measuring distances to stars, and it remains the only technique that tells us stellar distances without any assumptions about the nature of stars. If we know a star's distance from parallax, we can calculate its luminosity with the inverse square law for light. We now have parallax measurements for thousands of stars, which is a large enough number that astronomers have been able to draw some general conclusions about them. As we'll see later, these lessons have taught astronomers how to estimate luminosities for many more stars, even without knowing their distances. Astronomers today often use the inverse square law for light to calculate distances to objects for which we can reliably estimate luminosities, as well as to calculate luminosities of objects for which we have measured distances.

Parallax measurements have given us detailed knowledge of what our local solar neighborhood is like. For example, we know of more than 300 stars within about 33 light-years (10 parsecs) of the Sun. About half are binary star systems consisting of two orbiting stars, or multiple star systems containing three or more stars. Most are tiny, dim red stars—so dim that we cannot see them with the naked eye, despite the fact that they are relatively close. A few nearby stars, such as Sirius (8.6 light-years), Altair (17 light-years), Vega (25 light-years), and Fomalhaut (25 light-years), are white in color and bright in our sky, but most of the brightest stars in the sky lie farther away. Because so many nearby stars appear dim while many more distant stars appear bright, stellar luminosities must span a wide range.

The Luminosity Range of Stars Now that we have discussed how we determine stellar luminosities, it's time to take a quick look at the results; these results have been drawn both from stars for which we have parallax measurements and from those for which we determine distance in other ways. We usually state stellar luminosities in comparison to the Sun's luminosity, which we write as L_{Sun} for short. For example, Proxima Centauri, the nearest of the three stars in the Alpha Centauri system and hence the nearest star besides our Sun, is only about 0.0006 times as luminous as the Sun, or $0.0006L_{\text{Sun}}$. Betelgeuse, the bright left-shoulder star of Orion, has a luminosity of $38{,}000L_{\text{Sun}}$, meaning that it is

MATHEMATICAL INSIGHT 15.2

The Parallax Formula

We can derive the formula relating a star's distance and parallax angle by studying Figure 15.3. The parallax angle p is part of a right triangle, and from trigonometry you may recall that the *sine* of angle p is the length of the side opposite this angle divided by the length of the hypotenuse. Because the side opposite p is the Earth-Sun distance of 1 AU and the hypotenuse is the distance d to the object, we find

$$\sin p = \frac{\text{length of opposite side}}{\text{length of hypotenuse}} = \frac{1\ \text{AU}}{d}$$

Solving for d, the formula becomes

$$d = \frac{1\ \text{AU}}{\sin p}$$

By definition, 1 parsec is the distance to an object with a parallax angle of 1 arcsecond (1″), or $\frac{1}{3600}$ degree (because $1° = 60'$ and $1' = 60''$). Substituting these numbers into the parallax formula and using a calculator to find that $\sin 1'' = 4.84814 \times 10^{-6}$, we get

$$1\ \text{parsec} = \frac{1\ \text{AU}}{\sin 1''} = \frac{1\ \text{AU}}{4.84814 \times 10^{-6}} = 206{,}265\ \text{AU}$$

That is, 1 parsec = 206,265 AU. Converting units, we also find that 1 parsec $= 3.09 \times 10^{13}$ km = 3.26 light-years (because 1 AU = 149.6 million km and 1 light-year $= 9.46 \times 10^{12}$ km).

We need one more fact from geometry to derive the parallax formula given in the text. As long as the parallax angle, p, is small, $\sin p$ is proportional to p. For example, $\sin 2''$ is twice as large as $\sin 1''$, and $\sin \frac{1}{2}''$ is half as large as $\sin 1''$. (You can verify these examples with your calculator.) If we use $\frac{1}{2}''$ instead of 1″ for the parallax angle in the formula above, we get a distance of 2 parsecs instead of 1 parsec. Similarly, if we use a parallax angle of $\frac{1}{10}''$, we get a distance of 10 parsecs. Generalizing, we get the simple parallax formula given in the text:

$$d\ (\text{in parsecs}) = \frac{1}{p\ (\text{in arcseconds})}$$

EXAMPLE 1: Sirius, the brightest star in our night sky, has a measured parallax angle of 0.379″. How far away is Sirius in parsecs? In light-years?

SOLUTION:

Step 1 Understand: We are given the parallax angle for Sirius in arcseconds, so we use the parallax formula to find its distance. Because the parallax angle is between 0.1″ and 1″, we expect the answer to be a distance between 1 and 10 parsecs.

Step 2 Solve: Substituting the parallax angle of 0.379″ into the formula, we find that the distance to Sirius in parsecs is

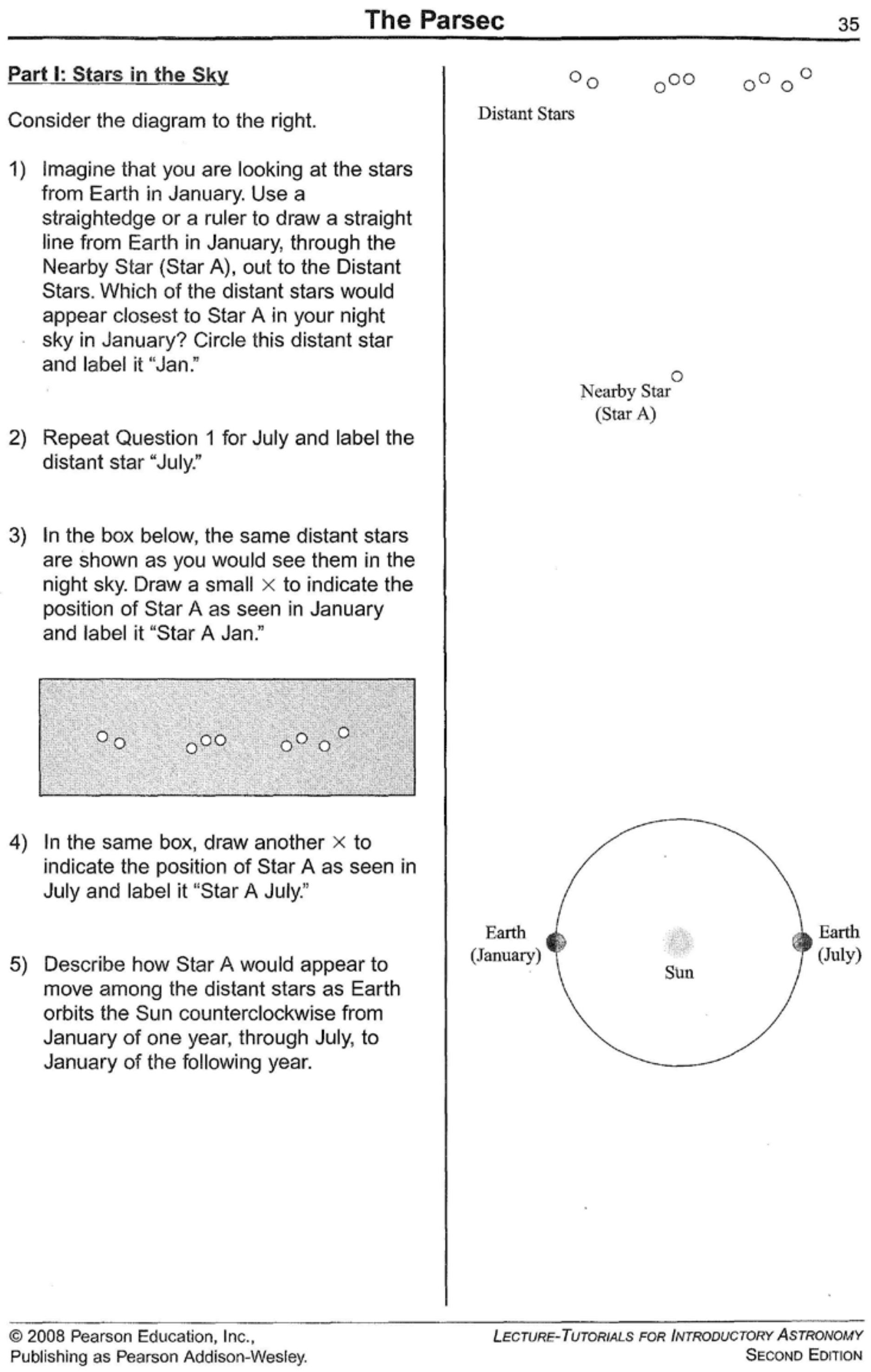
$$d\ (\text{in parsecs}) = \frac{1}{0.379} = 2.64\ \text{pc}$$

Because 1 parsec = 3.26 light-years, this distance is equivalent to

$$2.64\ \text{parsecs} \times 3.26\,\frac{\text{light-years}}{\text{parsec}} = 8.60\ \text{light-years}$$

Step 3 Explain: From its measured parallax angle, we have found that the distance to Sirius is 2.64 parsecs, or 8.60 light-years.

APPENDIX B



The Parsec

35

Part I: Stars in the Sky

Consider the diagram to the right.

1) Imagine that you are looking at the stars from Earth in January. Use a straightedge or a ruler to draw a straight line from Earth in January, through the Nearby Star (Star A), out to the Distant Stars. Which of the distant stars would appear closest to Star A in your night sky in January? Circle this distant star and label it "Jan."

2) Repeat Question 1 for July and label the distant star "July."

3) In the box below, the same distant stars are shown as you would see them in the night sky. Draw a small × to indicate the position of Star A as seen in January and label it "Star A Jan."

4) In the same box, draw another × to indicate the position of Star A as seen in July and label it "Star A July."

5) Describe how Star A would appear to move among the distant stars as Earth orbits the Sun counterclockwise from January of one year, through July, to January of the following year.

Lecture-Tutorials for Introductory Astronomy
Second Edition

The apparent motion of nearby objects relative to distant objects, which you just described, is called **parallax.**

6) Consider two stars (C and D) that both exhibit parallax. If Star C appears to move back and forth by a greater amount than Star D, which star do you think is actually closer to you? If you're not sure, just take a guess. We'll return to this question later in this activity.

Part II: What's a Parsec?

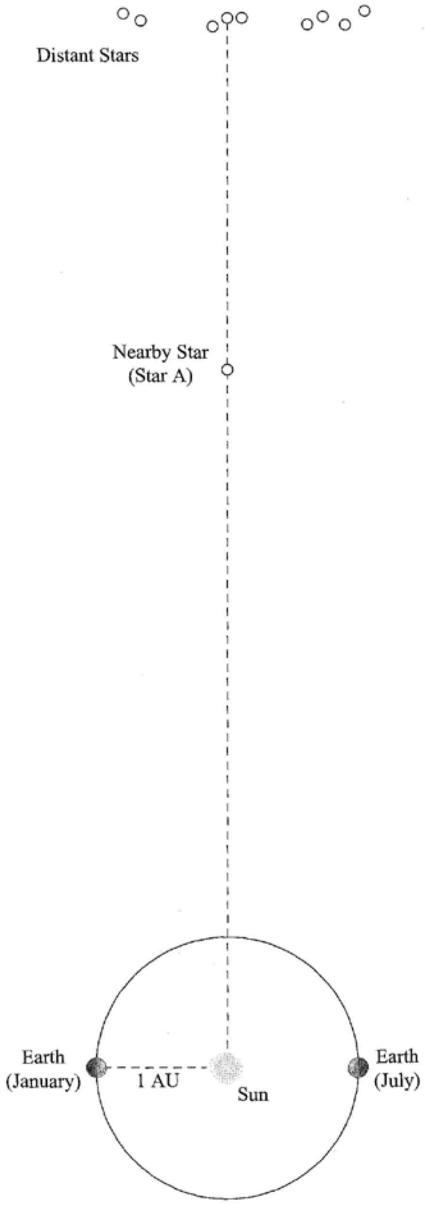
Consider the diagram to the right.

7) Starting from Earth in January, draw a line through Star A to the top of the page.

8) There is now a narrow triangle, created by the line you drew, the dotted line provided in the diagram, and the line connecting Earth and the Sun. The small angle, just below Star A, formed by the two longest sides of this triangle is called the **parallax angle** for Star A. Label this angle **"p_A."**

Knowing a star's parallax angle allows us to calculate the distance to the star. Since even the nearest stars are still very far away, parallax angles are extremely small. These parallax angles are measured in "arcseconds" where an arcsecond is 1/3600 of 1 degree.

To describe the distances to stars, astronomers use a unit of length called the **parsec.** One parsec is defined as the distance to a star that has a **par**allax angle of exactly 1 arc**sec**ond. The distance from the Sun to a star 1 parsec away is 206,265 times the Earth–Sun distance or 206,265 AU. (Note that the diagram to the right is not drawn to scale.)



9) If the parallax angle for Star A (**p_A**) is 1 arcsecond, what is the distance from the Sun to Star A? (Hint: Use parsec as your unit of distance.) Label this distance on the diagram.

10) Is a parsec a unit of length or a unit of angle? It can't be both.

Note: Since the distance from the Sun to even the closest star is so much greater than 1 AU, we can consider the distance from Earth to a star and the distance from the Sun to that star to be approximately equal.

Part III: Distances

11) Consider the following debate between two students regarding the relationship between parallax angle and the distance we measure to a star.

Student 1: *If the distance to the star is more than 1 parsec, then the parallax angle must be more than 1 arcsecond. So a star that is many parsecs away will have a large parallax angle.*

Student 2: *If we drew a diagram for a star that was much more than 1 parsec away from us, the triangle in the diagram would be pointier than the one we just drew in Part II. That should make the parallax angle smaller for a star farther away.*

Do you agree or disagree with either or both of the students? Explain your reasoning.

12) On your diagram from Part II, draw a second star along the dotted line farther from the Sun than Star A and label this far away star "Star B." Repeat steps 7 and 8 from Part II, except label the parallax angle for this Star B with **p_B**.

13) Which star, the closer one (Star A) or the farther one (Star B), has the larger parallax angle?

14) Check your answers to Questions 6 and 11 and resolve any discrepancies.

Part I: Angular Measurement

Imagine that you are standing in an open field. While facing south, you see a house in the distance. If you turn your head and look directly east (to your left), you see a barn in the distance.

1) What is the angle between you, the house, and the barn? (Hint: If you point at the barn with one arm and point at the house with your other arm, what angle do your arms make?)

2) You see the Moon on the horizon just above the barn in the east and also see a bright star directly overhead. What is the angle between you, the Moon, and the overhead star?

3) Compare your answers for the barn–house angle from Question 1 and the Moon–star angle from Question 2. Are they the same?

4) Do the angles from above tell you anything about the actual distance between the barn and house or the Moon and star?

We are often unable to **directly** measure distances to faraway objects in our night sky. However, we can obtain the distances to relatively nearby stars by using their parallax angles. Because even these stars are very far away (up to about 500 parsecs), the parallax angles for these stars are very small. They are measured in units of **arcseconds**, where 1 arcsecond is 1/3600 of 1 degree. To give you a sense of how small this angle is, the thin edge of a credit card, when viewed from one football field away, covers an angle of about 1 arcsecond.

Part II: Finding Stellar Distance Using Parallax

Consider the star field drawing shown in Figure 1. This represents a tiny patch of our night sky. In this drawing we will imagine that the angle separating Stars A and B is just 1/2 of an arcsecond.

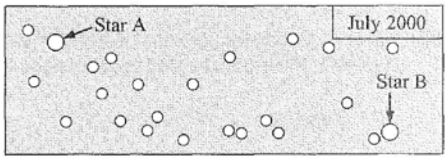


Figure 1

In Figure 2 (see the final page of the activity) there are drawings of this star field taken at different times during the year. One star in the field moves back and forth across the star field (exhibits parallax) with respect to the other, more distant stars.

5) Using Figure 2, determine which star exhibits parallax. Circle that star on each picture in Figure 2.

6) In Figure 1, draw a line that shows the range of motion for the star you saw exhibiting parallax in the drawings from Figure 2. Label the end points of this line with the months when the star appears at those end points.

7) How many times bigger is the separation between Stars A and B compared to the distance between the end points of the line showing the range of the motion for the star exhibiting parallax?

8) Recall that Stars A and B have an angular separation of 1/2 of an arcsecond in Figure 1. Consider two more stars (C and D) that are separated **twice** as much as Stars A and B. What is the angular separation between Stars C and D in arcseconds?

9) What is the angular separation between the end points that you marked in Figure 1 for the nearby star exhibiting parallax?

Note: We define a star's **parallax angle** as **half** the angular separation between the end points of the star's angular motion.

10) What is the parallax angle for the nearby star exhibiting parallax from Question 9?

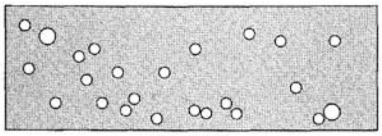
Note: We define 1 **parsec** as the distance to an object that has a **par**allax angle of 1 arc**sec**ond. For a star with a parallax angle of 2 arcseconds, the distance to the star from Earth would be 1/2 of a parsec.

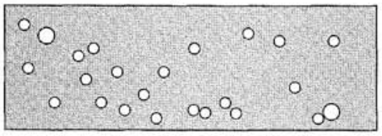
11) For a star with a parallax angle of 1/2 of an arcsecond, what is its distance from us?

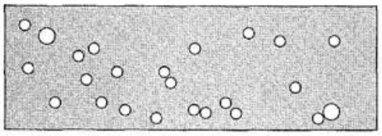
12) For a star with a parallax angle of 1/4 of an arcsecond, what is its distance from us?

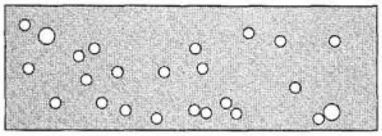
13) What is the distance from us to the nearby star exhibiting parallax in the drawings from Figure 2? (Hint: Consider your answer to Question 10.)

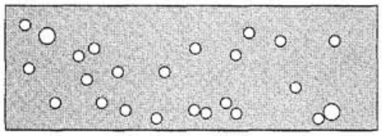
a) 1 parsec
b) 2 parsecs
c) 4 parsecs
d) 8 parsecs
e) 16 parsecs

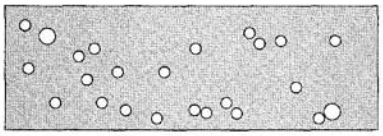

July 2000

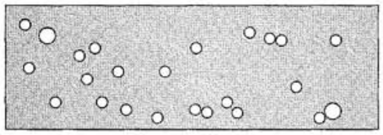

August 2000

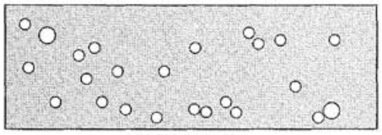

September 2000

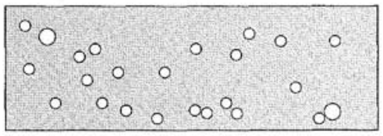

October 2000

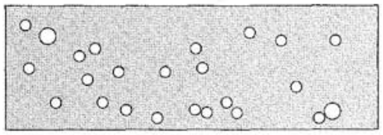

November 2000

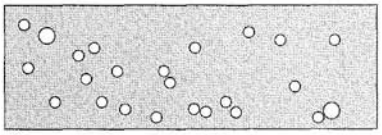

December 2000

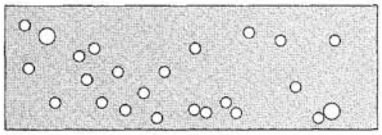

January 2001

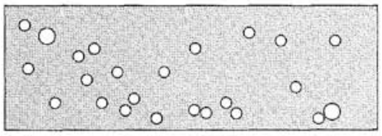

February 2001


March 2001


April 2001


May 2001


June 2001


July 2001

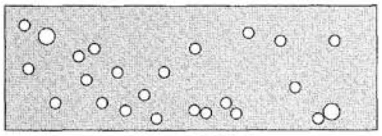

August 2001

Figure 2

Adapted from: Astronomy Diagnostic Test (Hufnagel et al., 1999), Concepts Tests Library (Green, 2003), and Well-Structured Problem-Solving Processes Inventory (WPSPI). Shin, Jonassen, & McGee 2003.

APPENDIX C

Stellar Parallax Assessment (*SPA*)

1. **Recall. Distance, Baseline, Parallax.** The distances to nearby stars may be measured by observing their apparent motion across the celestial sphere as
   a) the Earth orbits around the Sun.
   b) the Earth rotates on its axis.
   c) the Sun orbits around the center of the Galaxy.
   d) the planets cross their path.

2. **Recall. Parallax.** What is parallax in astronomy? (Explain your answer).
   ____
   ____

3. **Recall. Distance and Parallax angle.** Rank the order of parallax angle (from the greatest to the least) of the following stars:
   a) Spica (400 ly away from the sun)
   b) Acturus (300 ly away from the sun)
   c) Rigel
   d) Antares
   e) Polaris
   (greatest) [__], [__], [__], [__], [__] (least)

4. **Distance and Angular size.** If the Moon were twice as far away from the Earth as it is now, how would its appearance change?
   a) it would not change.
   b) it would look about twice as big.
   c) it would look about twice as small.
   d) it would look very small (more than twice as small).
   e) it would look very big (more than twice as big).

5. **Distance and Parallax.** You photograph a region of the night sky in March, in September, and again the following March. The two March photos look the same, but the September photo shows three stars in different locations. The star whose position shifts the most must be

   a) farthest away, because ____

   b) closest, because ____

   c) receding from Earth most rapidly, because ______________________________

   d) perpendicular to the plane of the solar system, because ____________________

   ____

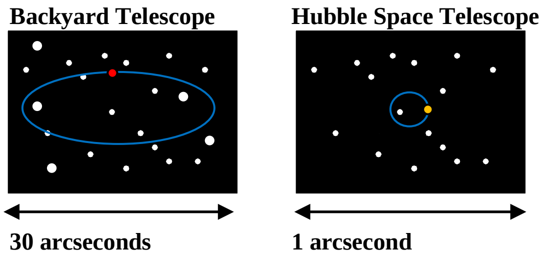
6. **Baseline and Parallax.** Two spacecrafts are exploring our solar system: one is near the planet Venus and another one is near Pluto. They both measured the parallax shift of a distant star. Which spaceship will measure the largest parallax shift?
   a) Spacecraft near Venus
   b) Spacecraft near Pluto
   c) The both would measure the identical shift
   d) Neither, since you have to be 1AU away from the sun to measure the parallax shift

7. **Distance and Angular Size.** From Earth, planet A spans an angle of 5 arcseconds, and planet B spans an angle of 10 arcseconds. If the radius of planet A equals the radius of planet B,
   a) planet A is twice as big as planet B.
   b) planet A is twice as far as planet B.
   c) planet A is half as far as planet B.
   d) planet A and planet B are the same distance.
   e) planet A is five times as far as planet B.

8. **Parallax Angle and Parallax.** Suppose you measure the parallax angle for a particular star to be 0.2 arcsecond. The parallax angle of his star is:
   a) 0.1 arcsecond
   b) 0.2 arcsecond
   c) 0.4 arcsecond
   d) 5 arcseconds

9. **Resolution.** If you cannot see any parallax shift with the equipment you have, then in order to observe the parallax you should
   a) Increase your equipment resolution so you would have more pixels per arcsecond
   b) Increase your equipment resolution so you would have more arcseconds per pixel
   c) You should give up this equipment because the star is too far away
   d) You should give up because the star is exactly in your line of sight

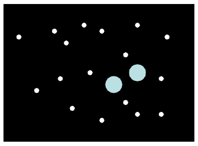
10. **Resolution and conversion.** Look very carefully at these two photographs. The both photographs show the parallax shift of a star labeled A and B which star is furthest away from us?



   a) Star A
   b) Star B
   c) We would need to know how many pixels in photographs before we can answer the question.
   d) We cannot tell because two photographs cover the different regions of the sky.

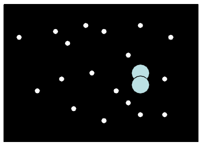
Explain your reasoning:

11. **Distance and Parallax.** Star A is at a distance of 10 lightyears away from the Earth, star B is only 5 lightyears away. Both stars appear close together on the celestial sphere, as shown in the picture.

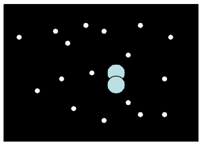


The picture does not say which star is A or B, but the two stars show up as fat dots.

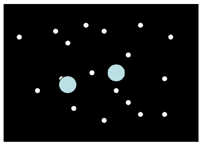
When we look at the stars 6 months later, which picture could correspond to what we see? Use the fixed stars in the picture [the small white dots] as your reference to determine the amount of movement.



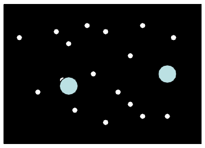
(A)



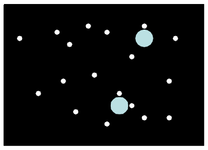
(B)



(C)



(D)



(E)

12. **Distance and Parallax angle.** Briefly explain why there are no stars except for the sun displaying the parallax angle of more than 1 arcseconds?

____

____

13. **Stellar Parallax.** Suppose that a new star was found which believed to be located some 100 lightyears away. What do you have to do to accurately determine its distance? List your procedural steps, include formula(s) (if any is required) and describe how these calculations will help you measure the distance to the star.

✓ ____

✓ ____

✓ ____

✓ ____

✓ ____

14. **Stellar Parallax angle and Distance.** Imagine that you are a science consultant for a science fiction movie. A plot of the movie is simple: a crew of explorers was traveling through the Milky Way galaxy when something went wrong. They woke up after a very long cryosleep to find out that instead of approaching Earth their spaceship was orbiting an unknown star somewhere in our galaxy. A spaceship's navigation system found the Sun and showed that a recorded parallax shift of the sun from the current ships orbit (diameter of 2 AU) is 0.5 arseconds. Is it a possible scenario? Now it is your turn to either show the actors how their characters should find their way back home or persuade a movie director that the scenario is scientifically unsound and from the given information the crew would never be able to find its way back to Earth.

____

____

____

____

____

APPENDIX D

Semi-structured Interview Questions

1. Do you use computers? How often do you use them? Do you own a computer (desktop/laptop)?

2. How comfortable are you with computers? Would you consider yourself a proficient computer user?

3. Do you play computer and/or video games? If you do, how often do you play? What is your favorite game? What do you have to do to win the game?

4. You are currently enrolled in an astronomy course. Have you taken any astronomy courses before? Was it in your high school? (If the answer is yes: Were you familiar with the concept of stellar parallax? Did you learn it in your high school's science class?)

5. Why are you taking the astronomy course right now? Are you taking it as a general education requirement or did you want to take it for other reasons?

6. What can you tell me about the stellar parallax? What is it? And, while you are explaining it to me, why not to use this piece of paper to draw a sketch.

7. Also, can you write down a formula for finding these distances to the stars? (If a student does not remember that the stellar parallax is used for finding distances to the stars: Do you remember a formula that comes along with the concept? Do you remember what is it used for?)

8. While you were sketching the concept of the stellar parallax and thinking about the formula, what were you referring to in your memories? What is the source of these memories?

9. Very good. Now let imagine a situation (a very hypothetical one) in which you observe two stars over a certain period of time and you notice (you are using a very good equipment for your observations, of course) that one of these stars appears to shift more while another shifts less. What can you tell me about these stars?

10. While you were learning the concept, you worked with *SPIRUT* (the computer-based tutorial), or *LTs* (the paper-based tutorial). How much do you remember about your experience with it? Does anything stand out in your memory?

11. Based on how much you remember today, how hard or easy was for you to work with the tutorial?

12. Were there any helpful features that you remember?

13. (*SPIRUT* only) Do you remember whether you had a full control over your actions in SPIRUT or you were guided throughout the tutorial?

14. (If restricted, the program-control option): Did you like that *SPIRUT* was restricting your actions and took you to different parts of the tutorial that it "thought" you had to brush up on or you would rather to be in control of the tutorial and chose what exercises you wanted to do, and in which order?

    (If unrestricted, the learner-control option): Did you like to be in control of your actions or would you prefer to let *SPIRUT* take the control and take you to different parts of the tutorial that it "thinks" you have to brush up on?

15. (Learner-controlled *SPIRUT* only) How did you use your freedom to control the environment?

16. Do you think adding gaming elements to the tutorial would make it more interesting for you to learn?

17. Did you work with a partner or independently?

18. (If worked with a partner): Did you like working with your partner? Can you describe to me how you worked on the exercises?

    (If worked independently): Did you like working by yourself or would you rather work with a partner?

19. Who do you think knew more about the concept: you or your partner?

REFERENCES

Adams, W., Reid, S., LeMaster, R., McKagan, S., Perkins, K., & Wieman, C. (2008). A study of educational simulations part I & II. *Journal of Interactive Learning Research*, (in press).

Adams, J., & Slater, T. (2000). Astronomy in the national science education standards. *Journal of Geoscience Education, 48*(1), 39-45.

Adams, W., Perkins, K., Podolefsky, N., Dubson, M., Finkelstein, N., & Wieman, C. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado learning attitudes about science survey. *Physical Review Special Topics - Physics Education Research, 2,* 1–14.

Aldrich, C. (2004). *Simulations and the future of learning*. San Francisco: Pfeiffer.

Anderson, J., Reder, L., & Simon, H. (1996). Situated learning and education. *Educational Researcher*, 25(4), 5-11.

Ardac, D., & Sezen, A., (2002). Effectiveness of computer-based chemistry instruction in enhancing the learning of content and variable control under guided versus unguided conditions. *Journal of Science Education and Technology, 11*(1),39–48.

Azevedo, R., Cromley, J. G., & Siebert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*, 344–370.

Bailey, J. (2004). A review of astronomy education research. *Astronomy Education Review*, *2*(2), 20-45.

Bailey, J., Prather, E., & Slater, T. (2004). Reflecting on the history of astronomy education research to plan for the future. *Advances in Space Research*, *34*(10), 2136-2144.

Bandura, A. (1977). *Social Learning Theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1986). *Social Foundations of Thought and Action*. Englewood Cliffs, NJ: Prentice-Hall.

Barab, S. (2003). An introduction to the special issue: Designing for virtual communities in the service of learning, *The Information Society*, *19*, 197–201.

Barab, S., & Dede, C. (2007). Games and immersive participatory simulations for science education: An emerging type of curricula. *Journal of Science Education and Technology*, *16*(1), 1-3.

Barab, S. & Duffy, T. (2000). From practice fields to communities of practice. In D. H. Jonassen, & S. M. Land (Eds.), *Theoretical foundations of learning environments*. pp. 25-55. Mahwah, NJ: Lawrence Erlbaum Associates.

Barab, S., Hay, K., Barnett, M., & Keating, T. (2000). Virtual solar system project: Building understanding through model building. *Journal of Research in Science Teaching*, *37*(7), 719-756.

Bell, B., & Kozlowski, S. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied Psychology*, *93*(2), 296-316.

Bennett J., Donahue. M., Schneider, N., & Voit, M. (2010). *The Cosmic Perspective.* 6th ed. San Francisco, CA: Pearson Addison-Wesley.

Berson, M. (1996). Effectiveness of computer technology in the social studies: A review of the literature. *Journal of Research on Computing in Education*, *28*(4), 486-499.

Bloom, B. (1956). *Taxonomy of educational objectives: The classification of educational goals. Handbook I: Cognitive Domain.* New York, NY, Longmans.

Blumenfeld, P., Soloway, E., Marx, R., Krajcik, J., Guzdial, M., & Palincsar, A. (1991). Motivating project-based learning: Sustaining the doing, supporting the learning. *Educational Psychologist*, *26*(3 & 4), 369-398.

Bodemer, D., Ploetzner, R., Feuerlein, I., & Spada, H. (2004). The active integration of information during learning with dynamic and interactive visualizations, *Learning and Instruction, 14,* 325-341.

Bostrom, R., Gupta, S., & Hill, J. (2008). Peer-to-peer technology in collaborative learning networks: applications and research issues. *International Journal of Knowledge and Learning*, *4*(1), 36-57.

Bransford, J., Brown, A., & Cocking, R. (2000). *How People Learn: Brain, Mind, Experience, and School*. Washington, D.C.:National Academy Press.

Briggs, D., Alonzo, A., Schwab, C., & Wilson, M. (2006). Diagnostic assessment with ordered multiple-choice items. *Educational Assessment*, *11*(1). 33-63.

Bryant, P., & Squire, S. (2001). Children's mathematics: Lost and found in space, in M. Gattis (ed.), *Spatial Schemes and Abstract Thought*, The MIT Press, Cambridge, pp. 175–200.

Carrie, C. (1984). Do learners make good choices? A review of research on learner control in institution. *Instructional Innovator*, *29*(2), 15-17.

Campbell, V. (1964). Self-direction and programmed instruction for five different types of learning objectives. *Psychology in the Schools*, *1*(4), 348-359.

Campbell, V., & Chapman, M. (1967). Learner control versus program control of instruction. *Psychology in the Schools*, *4*(2), 121-130.

Cameron, B. (2003). The Effectiveness of Simulation in a Hybrid and Online Networking Course. Techtrends, 47(5).

Chang, K-E., Sung, Y-T., & Lee, C-L. (2003). Web-based Collaborative inquiry learning. *Journal of Computer Assisted Learning*, *19,* 56-69.

Chen, C., Yang, J., Shen, S., & Jeng, M. (2007). A desktop virtual reality Earth motion system in astronomy education. *Educational Technology & Science*, *10*(3), 289-304.

Clements, D., & Sarama, J. (2004). Building Blocks for early childhood mathematics. *Early Childhood Research Quarterly, 19*, 181-189.

Cohen, J. (1994). The earth is round *(p* < .05). *American Psychologist*, *49*(12), 997-1003.

Colburn, A. (2000). An inquiry primer. *Science Scope, 23*, 42-44.

Collins, A. (1985). Teaching reasoning skills. In S. Chipman, J. Segal, & R. Glaser (eds.) *Thinking and Learning Skills, Vol. 2*. Hillsdale, NJ: Erlbaum.

Collins, A., Brown, I., & Newman, S. (1989). Cognitive apprenticeship: Teaching the craft of reading, writing and mathematics. In L. B. Resnick (Ed.), *Knowing, learning and instruction: Essays in honor of Robert Glaser*. Hillsdale, NJ: Erlbaum.

Crooks, S., Klein, J., Savenye, W., & Leader, L. (1998). Effects of cooperative and individual learning during learner-controlled computer-based instruction. *Journal of Experimental Education, 66*(3), 223–24.

Dede, C. (2005). Planning for neomillennial learning styles: implications for investments in faculty and technology. In D. Oblinger & J. Oblinger (Eds), *Educating the Net generation* (pp. 15.1–15.22). Boulder, CO: EDUCAUSE. Retrieved March 20, 2006, from http://www.educause.edu/educatingthenetgen

Dede, C., Salzman, M., Loftin, R. B., & Sprague, D. (1999). Multisensory immersion as a modeling environment for learning complex scientific concepts. In W. Feurzeig & N. Roberts (Eds.), *Modeling and simulation in science and mathematics education.* New York: Springer Verlag.

Dennen, V. (2002). Cognitive apprenticeship in educational practice: Research on scaffolding, modeling, mentoring, and coaching as instructional strategies. In D.H. Jonassen (Ed.), *Handbook of research on educational communications and technology (2nd ed.)* (pp. 571-581). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Dewey, J. (1916). *Essays in Experimental Logic.* Chicago, IL: The University of Chicago Press.

Dewey, J. (1938). *Logic: The Theory of Inquiry,* New York: Holt and Co.

Diakidoy, I., & Kendeou, P. (2001). Facilitating conceptual change in astronomy: A comparison of the effectiveness of two instructional approaches. *Learning and Instruction*, *11*, 1–20.

Dickey, M., (2005). Three-dimensional virtual worlds and distance learning: two case studies of Active Worlds as a medium for distance education, *British Journal of educational Technology*, 36(3), 439-451.

Duffy, T., & Cunningham D. (1996). *Constructivism: Implications for design and delivery of instruction*. Handbook of research for educational communications and technology. D. H. Jonassen, New York: Simon and Schuster: 170-198.

Duffy, T., Dueber, B., & Hawley, C. (1998). Critical thinking in a distributed environment: A pedagogical base for the design of conferencing systems. In C. Bonk & K. King (Eds.) *Electronic collaborators: Learner-centered technologies for literacy, apprenticeship, and discourse (pp. 51-78).* New Jersey: Lawrence Erlbaum Associates, Publishers. Edelson, 2000

Edelson, D. (2001). Learning-for-use: A framework for integrating content and process learning in the design of inquiry activities. *Journal of Research in Science Teaching, 38,* 355–385.

Edelson, D., Gordin, D., & Pea, R. (1999). Addressing the challenges of inquiry-based learning through technology and curriculum design. *The Journal of the Learning Sciences* 8: 391– 450.

Edelson, D., Pea, R., & Gomez, L. (1996). Constructivism in the collaboratory. In B. G. Wilson (ed.), *Constructivist Learning Environments: Case Studies in Instructional Design* (Englewood Cliffs, NJ: Educational Technology Publications), 151-164.

Ellermann, H., & Free, E. (1990). A subject-controlled environment for paired associate learning. *Journal of Computer-Based Instruction*, *17*(3), 97-102.

Field, A. (2005). *Discovering statistics using SPSS*. 2nd ed. London: Sage.

Fraknoi, A. (2001). Enrollments in Astronomy 101 courses. *Astronomy education Review*. *1*(1), 121-123.

Fraser, B. (1998). Science learning environments: Assessment, effects and determinants. In B.J. Fraser & K.G. Tobin (Eds.), *International handbook of science education* (pp.527-564). Dordrecht, The Netherlands: Kluwer Academic Publishers.

Fridereksen, R., & White, B. (2002). Conceptualizing and constructing linked models: Creating coherence in complex knowledge systems. In B. Brna (Ed.), *The role of communication in learning to model*, (pp. 69-96). Mahwah, NJ: Lawrence Erlbaum Associates.

Gay, G. (1986). Interaction of learner control and prior understanding in computer-assisted video instruction. *Journal of Educational Psychology*, *78*(3), 225-227.

Garrison, D. (1992). Critical thinking and self-directed learning in adult education: An analysis of responsibility and control issues. *Adult Education Quarterly*, *42*(3), 136-148.

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.

Ge, X., & Land, S. (2003). Scaffolding students' problem-solving processes in an ill-structured task using question prompts and peer interactions. *Educational Technology Research and Development, 51*(1), 21-38.

Gentry, J., Dickinson, J., Burns, A., McGinnis, L., & Park, J. (2006). The Role of Learning versus Performance Orientations When Reacting to Negative Outcomes in Simulation Games. *Developments in Business Simulation and Experiential Exercises*, *33*, pp. 79-84, Reprinted in the *Bernie Keys Library*, 10th ed. [Available from http://ABSEL.org].

Gibbs, G. (1988). *Learning by Doing: A Guide to Teaching and Learning Methods.* London, UK: Further Education Unit.

Glassman, M. (2001). Dewey and Vygotsky: Society, Experience, and Inquiry in Educational Practice. *Educational Researcher*, *30* (4), 3-14.

Graesser, A. & Person, N. (1994). Question asking during tutoring. *American Educational Research Journal, 31*, 104-137.

Granott, N., & Parziale, J. (2002). Microdevelopment: A process-oriented perspective for studying development and learning. In N. Granott & J. Parziale (Eds.), *Microdevelopment: Transition processes in development and learning* (pp. 1–28). Cambridge, England: Cambridge University Press.

Gredler, M. (2004). Games and simulations and their relationships to learning. In D.H. Jonassen (Ed.), *Handbook of research on educational communications and technology (2nd ed.)* (pp. 571-581). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Green, P. (2003). *Peer Instruction for Astronomy*. Upper Saddle River, NJ: Pearson Education Inc.

Gronlund, N. (1982). *Constructing achievement tests*. New York: Prentice-Hall Inc.

Guiller, J., Durndell, A., & Ross, A. (2007). Peer interaction and critical thinking: face-to-face or online discussion? *Learning and Instruction*, *18*, 187-200.

Gunawardena, C. (1995). Social presence theory and implications for interaction collaborative learning in computer conferences. *International Journal of Educational Telecommunications*, 1(2/3), 147–166.

Hannafin, M. (1984). Guidelines for using locus of instructional control in the design of computer-assisted instruction, *Journal of Instructional Development*, *7*(3), 6-10.

Hartley, J. (1985). Some psychological aspects of computer-assisted learning and teaching. *Programmed Learning and Educational Technology, 22*(2), 140-149

Hawkins, J., & Pea, R. (1987). Tools for Bridging the Cultures of Everyday and Scientific Thinking. *Journal of Research in Science Teaching, 24*(4), 291-307.

Heift, T. & Nicholson, D. (2001). Web delivery of adaptive and interactive language tutoring. *International Journal of Artificial Intelligence in Education*, *12*(4), 310-325.

Henri, F. (1992). Computer conferencing and content analysis. In A. Kaye (Ed.) *Collaborative Learning Through Computer Conferencing: The Najaden papers* (p. 117-136). Berlin: Springer-Verlag.

Herman, J. (1994). Evaluating the use of technology in school reform. In B. Means (Ed.), *Technology and education reform: The reality behind the promise* (pp.133-167), San Francisco: Jossey-Bass.

Hernández-Leo, D., Villasclaras-Fernández, E. , Asensio-Pérez, J., Dimitriadis, Y., Jorrín-Abellán, I., Ruiz-Requies, I., et al. (2006). COLLAGE: A collaborative Learning Design editor based on patterns. *Educational Technology & Society*, *9*(1), 58-71.

Hirshfeld, A. (2001). *Parallax: Race to Measure the Cosmos*. Henry Holt and Company: New York.

Hmelo, C. (1998). Problem-based learning: effects on the early acquisition of cognitive skill in medicine. *Journal of Learning Sciences*, *7*, 173–208.

Hmelo, C. (2001). Oncology thinking cap: Scaffolded use of a simulation to learn clinical trial design. *Teaching and Learning Medicine, 13*(3), 183-191.

Hmelo, C. & Day, R. (1999). Contextualized questioning to scaffold learning from simulations. *Computers and Education, 32*, 151-164.

Hmelo, C., & Guzdial, M. (1996). Of black and glass boxes: Scaffolding for learning and doing. In D. C. Edelson & E. A. Domeshek (Eds.), *Proceedings of ICLS 96* (pp. 128–134). Charlottesville, VA: AACE.

Hmelo-Silver, C. (2006). Design principles for scaffolding technology-based inquiry. In A. M. O'Donnell, C. E. Hmelo-Silver, & G. Erkens (Eds.), *Collaborative reasoning, learning and technology* (pp. 147–170). Mahwah, NJ: Erlbaum.

Hmelo-Silver, C., Duncan, R., & Chinn, C. (2007). Scaffolding and achievement in problembased and inquiry learning: A response to Kirschner, Sweller, and Clark. *Educational Psychologist, 42*(2), 99–107.

Hooper, S. (1992). Effects of peer interaction during computer-based mathematics instruction. *Journal of Educational Research, 85*(3), 180–189.

Hooper, S., Temiyakarn C. & Williams, M. (1991). The effects of cooperative learning and learner control on high- and average-ability students. *Educational Technology Research and Development*, *41*(2), 5-18.

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

Johansen & Tennyson, (1983). Effect of adaptive advisement on perception in learner controlled, computer-based instruction using rule-learning task. *Educational Communications and Technology Journal*, *31*(4), 226-236.

Jonassen, D. (1995). Computers as cognitive tools: Learning with technology not from technology. *Journal of Computing in Higher Education*, *6*(2) pp.40-73

Jonassen, D. (1996). *Computers in the classroom: Mindtools for critical thinking.* Englewood Cliffs, NJ: Prentice Hall.

Jonassen, D. (2000). *Computers as Mindtools for schools* (2nd ed.). Columbus OH: Merrill.

Jonassen, D., & Land, S. (2000). *Theoretical foundations of learning environments.* Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Jonassen, D., Prevish, T., Christy, D., & Stavurlaki, E. (1999). Learning to solve problems on the web: Aggregate planning in a business management course. *Distance Education: An International Journal, 20*(1), 49-63.

de Jong, T. (2006a). Computer simulations: Technological advances in inquiry learning. *Science, 312*, 532-533.

de Jong, T. (2006b). Scaffolds for computer simulation based scientific discovery learning. In J. Elen & R.E. Clark (Eds.), *Dealing with Complexity in Learning Environments*, pp. 107-128. Elsevier Science Publishers, London.

de Jong, T., & van Joolingen, W. (1998). Scientific discovery learning with computer simulations of conceptual domains. *Review of Educational Research*, *68,* 179-201.

van Joolingen, W., & de Jong, T. (1991). Characteristics of simulations for instructional settings. *Education and Computing*, *6*, 241-262.

van Joolingen, W., de Jong, T., & Dimitrakopoulout, A. (2007). Issues in computer supported inquiry learning in science. *Journal of Computer Assisted Learning*, *23*(2), 111-119.

Kasarskis, P., Stehwien, J., Hickx, J., Aretz A., & Wickens, C. (2001). Comparison of expert and novice scan behaviors during VFR flight. In *Proceedings of the 11th International Symposium on Aviation Psychology*. http://www.aviation.uiuc.edu/UnitsHFD/conference/proced01.pdf.

Kim, B., & Hay, K. (2004). *Creating universe with computer: Charting the structure of distribution between learner and a cognitive tool.* Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

King, A. (1998). Transactive peer tutoring: Distributing cognition and metacognition. *Educational Psychology Rev. 10*, 57-74.

King, A. (2002). Structuring peer interaction to promote high-level cognitive processing. *Theory into Practice 41*(1), 33-39.

Kirschner, P., Sweller, J., & Clark, R. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*, 75–86.

Kolodner, J. L., Owensby, J. N., & Guzdial, M. (2004). Case-based learning aids. In D.H. Jonassen (Ed.), *Handbook of research on educational communications and technology: A project of the Association for Educational Communications and Technology* (2nd ed., pp. 829-861). Mahwah, NJ: Lawrence Erlbaum Associates.

Kopcha, T., & Sullivan, H. (2007). Self-presentation bias in surveys of teachers' educational technology practices. *Educational Technology Research and Development, 56*(3), 265–286.

Krajcik, J. S., & Blumenfeld, P. (2006). Project-based learning. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 317–334). New York: Cambridge.

Kuhn, D., Black, J., Keselman, A., & Kaplan, D. (2000). The development of cognitive skills to support inquiry learning. *Cognition and Instruction, 18*, 495–523.

Lakkala, M., Lallimo, J., & Hakkarainenvan, K. (2005). Teachers' pedagogical designs for technology-supported collective inquiry: A national case study. *Computers & Education*, *45*, 337–356.

LaPointe, D., & Gunawardena, C. (2004). Developing, testing, and refining a model to understand the relationship between peer interaction and learning outcomes in computer-mediated conferencing. *Distance Education*, 25(1), 83–106.

Lave, J. (1988). *Cognition in practice.* New York: Cambridge University Press.

Lawless, K., & Brown, S. (1997). Multimedia learning environments: Issues of learner control and navigation. *Instructional Science*, *25*, 117-131.

Lenski, S., & Nierstheimer, S. (2002). Strategy instruction from a sociocognitive perspective. *Reading Psychology*, 23(2), 127–143.

Lepper, M., & Chabay, R. (1985). Intrinsic motivation and instruction: Conflicting views on the role of motivational process in computer-based education. *Educational Psychologist*, *20*(4), 217-230.

Lightman, A., & Sadler, P. (1993).Teacher predictions versus actual student gains. *The Physics Teacher*, *31*, 162–167.

Lim, C., Nonis D., & Hedberg, J. (2006). Gaming in a 3D multiuser virtual envoronmnet: engaging students in science lessons. *British Journal of Educational Technology*, 37(2), (pp. 211-231).

Lin, X., & Lehman, J., (1999). Supporting learning of variable control in a computer-based biology environment: effects of prompting college students to reflect on their own thinking. *Journal of Research in Science Teaching. 36*(7), 837-858.

Lyle. J. (2003). Stimulated Recall: a report on its use in naturalistic research. *British Educational Research Journal, 29*(6), 861-878.

Mager, R., & Clark, C. (1963). Explorations in student-controlled instruction. *Psychological Reports*, *13*(1), 71-76.

Manlove, S., & Lazonder, A. (2004). Self-regulation and collaboration in a discovery learning Environment, *Proceedings of First EARLI-SIG on Metacognition*, Amsterdam, The Netherlands.

Manlove, S., Lazonder, A., & de Jong, T. (2007). Software scaffolds to promote regulation during scientific inquiry to promote regulation during scientific Inquiry learning. *Metacognition & Learning*, (2), 141-155.

Marschall, A. (1998). Bringing the Universe into the Laboratory-Project CLEA: Contemporary Laboratory Exercises in Astronomy. In L. Gouguenheim, D. McNally, & J. R. Percy (Eds.), *New Trends in Astronomy Teaching*. Cambridge: Cambridge University Press.

Marschall, A. (2000). The Universe on a desktop: observational astronomy simulations in the instructional laboratory. *Astronomical Society of Australia*, *17*, 129-132.

Marschall, A., Snyder, A., & Cooper, P. (2000). A desktop universe for the introductory astronomy laboratory. *The Physics Teacher*, *38*, 536-537**.**

Mayer, E., Mautone, P., & Prothero, W., (2002). Pictorial aids for learning by doing in a multimedia geology simulation game. *Journal of Educational Psychology, 94*(1), 171-185.

Mayer, R., & Moreno, R. (2003). Nine Ways to Reduce Cognitive Load in Multimedia Learning. *Educational Psychologist, 38*(1), 43–52.

Mayer, R., & Sims, V. (1994). For whom is a picture worth a thousand words? Extensions of a dual-coding theory of multimedia learning. *Journal of Educational Psychology, 84*, 389–460.

Mazur, E. (1997). *Peer Instruction: A User's Manual*. Prentice-Hall, Upper Saddle River, NJ, 1997.

McNeil, B., & Nelson, K. (1991). Meta-analysis of interactive video instruction: A 10 year review of achievement effects. *Journal of Computer-Based Instruction, 18*(1), 1-6.

Milheim, W. (1990). The effects of pacing and sequence control in an interactive video lesson. *Educational Training Technology International*. 27(1), 7-19.

Mitrovic, A. (2003). An Intelligent SQL Tutor on the Web. *International Journal of Artificial Intelligence in Education, 13*(2-4), 171-195.

National Research Council. (1996). National Science Education Standards. Washington, D.C.: National Academy Press.

Neuman, W. (1997). *Social Research Methods: Qualitative and Quantitative Approaches*. Boston: Allyn and Bacon.

Newman, G., Johnson, C., Webb, B. & Cochrane, C. (1997). Evaluating the quality of learning in computer supported co-operative learning. *Journal of the American Society for Information Science, 48(6)*, 484-495.

Newton Suter, W. (2006). *Introduction to Educational Research: A Critical Thinking Approach.* Thousand Oaks, Calif.: SAGE Publications.

Nickerson, J., Corter, J., Esche, S., & Chassapis, C. (2007). A model for evaluating the effectiveness of remote engineering laboratories and simulations in education. *Computers & Education*, *49*(3), 708-725.

O'Rourke, L., Hatcher, L & Stepanski, E. (2005). *A Step-by-step approach to using SAS for univariate and multivariate statistics*. SAS Institute Inc. and John Wiley & Sons, Inc. North Carolina: Cary.

Orrill, C. (2001). Building technology-based, learner-centered classrooms: The evolution of a professional development framework. *Educational Technology Research and Development, 49*(1), 15-34.

Owensby, J., & Kolodner, J. (2002). Case interpretation and application in support of scientific reasoning. *Proceedings of Cognitive Science'2004,* Chicago 2004, 107-113.

Ozmen, H. (2007). The influence of computer-assisted instruction on students' conceptual understanding of chemical bonding and attitude toward chemistry: A case for Turkey. *Computers & Education*, *51*, 423-438.

Palincsar, A., Brown, A., & Martin, S. (1987). Peer interaction in reading comprehension instruction. *Educational Psychologist, 22*, 231-253.

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

Paivio, A. (1986). Mental Representations: A Dual Coding Approach. New York: Oxford University Press.

Paivio, A. (2006, September). Dual coding theory and education. Paper presented at the Pathways to Literacy Achievement for High Poverty Children, Ann Arbor, MI. Retrieved October 17, 2008, from http://www.umich.edu/~rdytolrn/pathways conference/presentations/paivio.pdf.

Pea, R. (1985). Beyond amplification: using the computer to reorganize mental functioning. *Educational Psychologist, 20*, 167-182.

Pellegrino, J., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know. Washington*, DC: National Academy Press.

Penland, P. (1979). Self-initiated learning. *Adult Education*, *29*(3), 170-179.

Pilkington, R., & Parker-Jones, C. (1996). Interacting with computer-based simulation: the role of dialog. *Computes and Education, 27*(1), 1-14.

Pollock, J., & Sullivan, H. (1990). Practice mode and learner control in computer-based instruction. *Contemporary Educational Psychology*, *15*(3), 251-260.

Postlethwait, S., Novak, J., & Murray, H. (1972). *The Audio-tutorial Approach to Learning*. 3rd ed. Burgess; Minneapolis, MN.

Prather, E., Slater, T., Adams, J., Bailey, J., Jones, L., & Dostal. J. (2004). Research on a lecture-tutorial approach to teaching introductory astronomy for non–science majors, *Astronomy Education Review 3*(2), 122.

Prather, E., Slater, T., Adams, J., & Brissenden, G. (2008). *Lecture-Tutorials for Introductory Astronomy*. 2nd ed. Sun Francisco, CA: Pearson Addison-Wesley.

Prather, E., Rudolph, A., Brissenden, G., & Schlingman, W. (2009). A national study assessing the teaching and learning of introductory astronomy. Part I. The effect of interactive instruction. *American Journal of Physics*, *77*(4) 320-330.

Purcell, K. & Andre, A. (2001) The Influence of callouts on pilot visual attention to an electronic taxi map. *Focusing Attention on Aviation Safety report*.

Rao, S., & DiCarlo, S. (2000). Peer instruction improves performance on quizzes. *Advances in Physiology Education, 24*, 51–55.

Reigeluth, C., & Stein, F. (1983). The elaboration theory of instruction. In C.M. Reigeluth (Ed.), *Instructional-design theories and models* (pp. 335-382). Hillsdale, NJ: Lawrence Erlbaum.

Reinking, D., & Schreiner, R. (1985). The effects of computer-mediated text on measures of reading comprehension and reading behavior. *Reading Research Quarterly*, *22*(5), 536-552.

Rieber, L. (1996). Animation as feedback in a computer-based simulation: Representations matter. *Educational Technology Research & Development*, *44*, 5-23.

Rieber, L. (1990). Using computer animated graphics in science instruction with children. *Journal of Educational Psychology*. *82*(1), 135-140.

Rieber, L. (2005). Multimedia Learning in Games, Simulations, and Microworlds. In R.E. Mayer (Ed.). *The Cambridge handbook of multimedia learning*, (pp. 549-567), NY: Cambridge University Press.

Rieber, L., Tzeng, S., Tribble, K., & Chu, G. (1996). Feedback and elaboration within a computer-based simulation: A dual coding perspective. 18th National Convention of the Association for Educational Communications and Technology, Indianapolis, IN.

Rieber,L., & Noah, D. (1997). Effect of gaming and graphical metaphors on reflective cognition within computer-based simulations. *Paper presented at the annual meeting of the American Educational Research Association*, Chicago, March.

Rieber,L., Noah, D., & Nolan, M. (1998). Metaphors as graphical representations within open-ended computer-based simulations. *Paper presented at the annual meeting of the American Educational Research Association*, San Diego, April.

Rieber, L., Tzeng, S., & Tribble, K. (2004). Discovery learning, representation, and explanation within a computer-based simulation: Finding the right mix. *Learning and Instruction*, 14, 307-323.

Romiszowski, A., & Mason, R. (1996). Computer-mediated communication. In D. Jonassen (Eds.), *Handbook of research for educational communications and technology* (pp. 438 - 456). New York: Macmillan.

Rooze, G. E. & Northup, T. (1989). *Computers, thinking, and social studies*. Englewood, CO: Teacher Ideas Press.

Ross, S., & Morrison, G. (2001). *Getting started in instructional technology research* (3rd ed.). Bloomington, IN: Association for Educational Communication and Technology. https://www.aect.org/intranet/Publications/Research/index.html.

Ross, S., & Rakow, E., (1981). Learner control versus program control as adaptive strategies for selection of instructional support on math rules. *Journal of Educational Psychology, 73*, 745-753.

Rus, V, Cai, Z, & Graesser, AC. (2007). Experiments on Generating Questions About Facts. In *Proc. CICLing*.

Russell, T., & Munby, H. (1991). Reframing: The role of experience in developing teachers' professional knowledge. In D. Schön (Ed.), *The reflective turn: Case studies in and on educational practice* (pp164-187). New York: Columbia University and the Teachers College Press.

Rutherford, A., (2001). *Introducing ANOVA and ANCOVA: A GLM Approach*. Sage, London.

Ruzhitskaya, L. & Speck, A. (2007). Finding Stellar Properties: From Parallax to Radius. *Cosmos in the Classroom 2007*, 2007, Astronomical Society of the Pacific, Claremont, CA. 139-142.

Ruzhitskaya, L., & Speck, A. (2008). Project CLEA - The Moons of Jupiter: Understanding the Kepler's Laws in Astronomy 101, *Proceedings of American Astronomical Society, AAS Meeting 212*, 40.02, Saint Louis.

Sadler, P. (1987). Misconceptions in astronomy. *Proceedings of the Second International Seminar on Misconceptions and Educational Strategies in Science and Mathematics*, Cornell University, Ithaca, NY, 422-425.

Sadler, P. (1996). Astronomy's Conceptual Heirarchy. In J.A. Percy (Ed.), Proceedings of an ASP symposium: *Astronomy education: current developments, future coordination*, 24-25 June 1994 (pp 46-60). College Park, MD: ASP.

Sadler, P. (2000). The Relevance of Multiple-Choice Tests in Assessing Science Understanding. Chap. 11. In: *Assessing Science Understanding: A Human Constructivist View*. (Joel J. Mintzes; James H.Wandersee; and Joseph, D. Novak, eds.) (Series ed.: Gary D. Phye. Educational Psychology Series.) San Diego, CA: Academic Press, 249–278.

Sadoski, M., & Paivio, A. (2001). *Imagery and text: A Dual Coding theory of Reading and Writing*. Mahwah, NJ: Lawrence Erlbaum Associates.

Saito H., & Miwa, K. (2007). Construction of a learning environment supporting learners' reflection: A case of information seeking on the web. *Computers and Education, 49*(2), 214-29.

Schwarz, B., & Linchevski, L. (2007). The role of task design and argumentation in cognitive development during peer interaction: The case of proportional reasoning. *Learning and Instruction*, *17*(5), 510-531.

Schank, R. (1999). *Dynamic memory revisited*. Cambridge, United Kingdom: Cambridge University Press.

Shank, Gary (2002). *Qualitative research: A personal Skills Approach.* Upper Saddle River, NJ: Merrill Prentice Hall.

Schank, R., Berman, T., & Macperson, K. (1999). Learning by doing. In C.M. Reigeluth (Ed.), *Instructional design theories and models: A new paradigm of instructional theory* (Vol. II) (pp. 161–181). Mahwah, NJ: Lawrence Erlbaum Associates.

Shin, N., Jonassen, D, & McGee, S. (2003). Predictors of well-structured and ill-structured problem solving in an astronomy simulation. *Journal of Research in Science teaching, 40*(1), 6-33.

Singhanayok, C., & Hooper, S. (1998). The effects of cooperative learning and learner control on students' achievement, option selections, and attitudes. *Educational Technology Research and Development*, *46*(2), 17-33.

Slavin, R. (1984). *Research Methods in Education: A Practical Guide*. Englewood Cliffs, N.J.: Prentice-Hall

Snow, R. (1980). *Aptitude, learner control, and adaptive instruction*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, CA. (ERIC Document Reproduction Service No. ED 180 447).

Solomon, G. (1983). The differential investment of mental effort in learning from different sources. *Educational Psychologist*, *18*(1), 42-50.

Solomon, G. (1984). Television is ‘‘easy’’ and print is ‘‘tough’’: The differential investment of mental effort in learning as a function of perceptions and attributions. *Journal of Educational Psychology*, *76*, 647–658.

Solomon, G., Globerson, T., & Guterman, E. (1989). The computer as a zone of proximal development: Internalizing reading-related metacognitions from a Reading Partner. *Journal of Educational Psychology*, *81*, 620-627.

Stieff M., & Wilensky, U. (2003). Connected chemistry – incorporating interactive simulations into the chemistry classroom. *Journal of Science Education and Technology*, *12*(3), 285-302.

Suh, M., & Moeyr-Packenham, P. (2007). The application of dual coding theory in multi-representational virtual mathematics environments. In J. Woo, H. Lee, K. Park, & D. Seo (Eds.), *Proceedings of the 31st Conference of the International Group for the Psychology of Mathematics Education*, 4, (pp.209-216). Seoul: PME.

Swinbank, W. (1997). Developing resources for astrophysics at A-level: the TRUMP Astrophysics Project. *Physics Education*, *32*(1), 40–45.

Taasoobshirazi, G., Zuiker, S., Anderson, K., & Hickey, D. (2006). Enhancing inquiry, understanding, and achievement in an astronomy multimedia learning environment. *Journal of Science Education and Technology*, *15*(5), 383-395.

Trochim, W. (2000). *The Research Methods Knowledge Base*. 2nd ed. Atomic Dog Publishing, Cincinnati, OH.

Vosniadou, S., & Brewer, W. (1992). Mental models of the earth: A study of conceptual change in childhood, *Cognitive Psychology*, *2*(4), 535-585.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Cambridge, MA: Harvard University Press.

Vygotsky, L. (1934/1994). The development of academic concepts in school aged children. In R. van der Veer & J. Valsiner (Eds.). The Vygotsky reader (355-370). Oxford: Blackwell. (Originally published in Russian 1934)

Warnakulasooriya, R. & Pritchard, D. (2005). Learning and problem-solving transfer between physics problems using Web-based homework tutor. In P. Kommers, & G. Richards, (Eds.), *Proceedings of World Conference on Educational Multimedia, Hypermedia, and Telecommunications 2005* (pp. 2976–2983). Chesapeake, VA: AACE.

Webb, N. (1982). Group composition, group interaction and achievement in cooperative small groups. *Journal of Educational Psychology, 74*, 475–484.

Wenger, E. (1998). *Communities of Practice: Learning, Meaning and Identity.* New York: Cambridge University Press.

White, B., & Frederiksen, J. (1998). Inquiry, modeling, and metacognition: Making science accessible to all students. *Cognition and Instruction* 16: 3–118. Windschitl, M., Winn, W., and Headley, N.

White, B. (1984). Designing computer games to help physics students understanding Newton’s laws of motion. *Cognition & Instruction*, *1*, 69-108.

Wieman, C., Perkins, K., & Adams, W. (2008). Oersted Medal Lecture 2007: Interactive simulations for teaching physics: What works, what doesn’t and why, *American. Journal of Physics. 76*(4-5), 393-399.

Williams, D. (1993, January). *A comprehensive review of learner-control: The role of learner characteristics*. Paper presented at the Convention of the Association for Educational Communications and Technology Sponsored by the Research and Theory Division, New Orleans, LA. (ERIC Document Reproduction Service No. ED 362 211)

Wilson, J., & Redish, E. (1989). Using computers in teaching physics. *Physics Today***,** *42*(1), 34-41.

Yair, Y., Schur, Y., & Mintz, R. (2003). A “Thinking Journey” to the planets using scientific visualization technologies: Implications to astronomy education. *Journal of Science Education and Technology*, *12*(1), 43-49.

Young, J. (1996). The effect of self-regulated learning strategies on performance in learner controlled computer-based instruction. *Educational Technology Research and Development*, *44*(2), 17-27.

Zacharia, Z., & Anderson, O. (2003). The effects of an interactive computer-based simulation prior to performing a laboratory inquiry-based experiment on students' conceptual understanding of physics. *American Journal of Physics*, *71*(6), 618-629.

Zelilik, M., Schau, C., & Mattern, N. (1998). Misconceptions and their change in university level astronomy courses. *The Physics Teacher, 36*(2), 104-107.

VITA

Lanika Ruzhitskaya was born in Russia. She attended Rostov-on-Don State University where she received her Bachelor's degree in Journalism. She received her M.A. in Advertising Management from the School of Journalism at the University of Missouri - Columbia (*MU*) and her Educational Specialist degree in Network Learning from the School of Information Science and Learning Technologies (*SISLT*) at *MU*. At *MU*, she received her Ph.D. in Information Science and Learning Technologies with emphasis area in Astrophysics.

Her research field is Astronomy Education in which she is developing and investigating the efficacy of computer-based learning technologies, as well as the application of the learning methods of inquiry-based learning, causality and scientific reasoning in astronomy and physics courses. She is studying the effectiveness of delivery of these methods in face-to-face, blended, and online learning settings.