GRAVITY & EINSTEIN: ASSESSING THE RUBBER SHEET ANALOGY IN UNDERGRADUATE CONCEPTUAL PHYSICS

by

Tiffany Rae Watkins

A thesis
submitted in partial fulfillment
of the requirements for the degree of
Master of Science in S.T.E.M. Education
Boise State University

August 2014

BOISE STATE UNIVERSITY GRADUATE COLLEGE

DEFENSE COMMITTEE AND FINAL READING APPROVALS

of the thesis submitted by

Tiffany Rae Watkins

Thesis Title:	Gravity & Einstein: Assessing the Rubber Sheet Analogy in Undergraduate Conceptual Physics
Date of Final Oral Examination:	13 June 2014

The following individuals read and discussed the thesis submitted by student Tiffany Rae Watkins, and they evaluated her presentation and response to questions during the final oral examination. They found that the student passed the final oral examination.

Laurie O. Cavey, Ph.D.	Chair, Supervisory Committee
Daryl J. Macomb, Ph.D.	Member, Supervisory Committee
Louis S. Nadelson, Ph.D.	Member, Supervisory Committee

The final reading approval of the thesis was granted by Laurie O. Cavey, Ph.D., Chair of the Supervisory Committee. The thesis was approved for the Graduate College by John R. Pelton, Ph.D., Dean of the Graduate College.

ACKNOWLEDGEMENTS

Completion of this thesis would not be possible without the continued encouragement and support of many people in my life...you know who you are.

ABSTRACT

The rubber sheet analogy for Einstein's General Relativity model of gravity is a popular way to visualize the effect mass has on the curvature of spacetime. My single group, quasi-experimental study with repeated measures was designed to assess the effectiveness of the rubber sheet analogy in teaching gravitational fields. I developed instructional materials, including a hands-on lab, to engage university students in thinking about gravity using the rubber sheet analogy. Previous research on students' ideas about gravity informed the development of the pre/post-test. My work is an important first step in establishing a standard assessment on gravity.

Approximately 97 students in a university-level conceptual physics course participated. The results are promising. Normalized gains for students were about 30%. Post-test scores indicate improvement in student recognition of factors that do and do not influence gravity. There were significant differences in student performance for two demographic categories (sex and age).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER ONE: INTRODUCTION

The intent of my study was to assess the effectiveness of the rubber sheet analogy and its ability to affect student understanding of gravitational fields. Einstein's Theory of General Relativity describes gravitational forces as the curvature of four-dimensional spacetime due to the presence of matter (Hartle, 2003). The rubber sheet analogy described below is a way to visualize the effect mass has on the curvature of spacetime.

Imagine a rubber sheet stretched out. Then, visualize placing a bowling ball on the sheet. The material around the bowling ball is stressed and deforms to the presence of the mass. Now take a golf ball and roll it across the sheet. The motion of the golf ball changes in response to the deformation of the sheet caused by the bowling ball (see Figure 1). The motion of the golf ball in the analogy is similar to how light passing near the sun, whose mass changes the light's trajectory. If the speed of the golf ball is fast enough, it continues to move by, on an altered course; but if the speed is too slow, the golf ball will fall in towards the bowling ball.

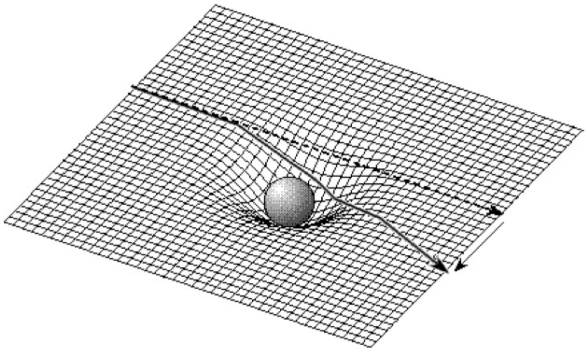


Figure 1. Scientific Diagram of Spacetime Curvature (Source:mattmancini.wordpress.com)

Examining how bodies move on the rubber sheet around masses can potentially be used by educators to initiate class discussions about orbiting bodies and in particular the effect of gravity. For example, a class discussion and demonstration may support student understanding of how satellites stay in orbit. Depending on the level of course, educators could also use the rubber sheet analogy to introduce the equations and calculations that allow for orbital speed determinations. I posit the rubber sheet analogy provides students with a mental model that supports thinking about and calculating gravity using Newton's Universal Law of Gravity.

The curvature of the sheet is representative of the familiar "g" constant physics students are introduced to in their first semester physics course. This "g" constant describes the acceleration an object experiences when falling near the surface in Earth's gravitational field. In general, the more massive the object, the greater the curvature, and thus, the larger the "g" constant for a given distance from the object. An educator can extend the rubber sheet analogy further by then having students imagine placing a less massive ball on the sheet to represent the moon. Even though this ball is less massive, the sheet deforms around this small object as well, albeit less so. The extension of the analogy may enable students to surmise that the "g" constant for the moon is smaller than Earth's acceleration constant because there is a weaker gravitational field. By doing so, the rubber sheet analogy could address the misconception that there is no gravity on the moon (Bar, Zinn, & Rubin, 1997; Borun, Massey & Lutter, 1993; Palmer, 2001).

The rubber sheet analogy could also provide educators with another method for presenting gravity besides the usual Newtonian, action-at-a distance, method of explanation. The possible benefit comes from providing a way for students to visualize why objects move as they do near massive objects. Action-at-a-distance describes the gravitational force, as well as electromagnetism, as being propagated by no physical connection between the two bodies. One of the problems with this method is that, due to our everyday experiences, many people think that a contact is needed for a force to be propagated (Watts, 1982). Students exposed to the visualization afforded by the rubber sheet analogy may develop a more comprehensive understanding of gravity (Bradamante & Viennot, 2007). With appropriate support, students may then be able to extrapolate the analogy to electrostatics and to what the electric field looks like around a charge.

There are many different methods researchers use to classify student understanding about gravity, mass, and weight. A review of research carried out between 1975 and 1985, Martinez (2001) summarized the methods of classification as Aristotelian, Newtonian, or some mixture of the two. Although the terminology used in the studies vary, the intent of the studies has been the same: to understand how students think about gravity.

I found only one study that reported the effectiveness of the rubber sheet analogy for teaching concepts of gravity (Baldy, 2007). Baldy’s (2007) study was conducted with 9th grade French students (*n*=123) and examined the effectiveness of two teaching methods: classical Newtonian (*n*=21) and a new approach using General Relativity and the rubber sheet analogy (*n*=102). My study builds on Baldy’s research by examining the conceptual understanding of American college students who were enrolled in a conceptual physics course where the rubber sheet analogy was used to teach ideas about gravity.

Research Questions

The documentation of student conceptions regarding gravitational fields is vast (Bar, et al., 1997; Borun, et al., 1993; Clement, Brown & Zietsman, 1989; Driver, 1989; Dostal, 2005; Galili & Bar, 1997; Galili, 2001; Graham & Berry, 1993; Watts, 1982). Some of these well-documented alternative conceptions (e.g. no gravity on the moon, gravity requiring an atmosphere, that gravity is associated with magnetism and rotation) could potentially be alleviated by using the rubber sheet analogy. In short, the rubber sheet analogy could provide a means for students to develop a robust understanding of gravity.

I have used the rubber sheet analogy in my teaching for many years and have always thought that the analogy could be very powerful for some students, especially those who need visual aids to help in understanding. Additionally, the rubber sheet analogy allows me to leverage student's prior experience with the analogy when we study electric fields and electric potential. Lastly, the analogy has seemed to correct some alternative conceptions that students might have. I wanted to create a lab, instructional materials, and instrument that could be used to explore this possibility.

The following research questions guided my investigation:

1. Is the rubber sheet analogy an effective model for teaching gravity?
2. What are student conceptions of gravity pre and post instruction?
3. Do personal characteristics or educational backgrounds appear to influence student conceptions?

My study could stimulate other educators to use the rubber sheet analogy in the classroom. Other researchers may be motivated to perform similar studies that investigate the effectiveness of the rubber sheet analogy, allowing for greater generalizability.

CHAPTER TWO: REVIEW OF THE LITERATURE

In this chapter, I describe results from educational research relevant to my study, in particular, students' understanding of gravity. I also address how research on conceptual understanding informs my perspective of learning, in general. Also included is a brief description of how the scientific community's perspective on gravity has developed over time, which is important given previous categorizations of student reasoning. Relativity is no longer something that only physicists can claim to understand, as the overall idea is accessible to the general population and allows for a deeper understanding of both spacetime and gravity, in particular.

Research into Conceptual Understanding

According to the tenets of constructivism, students come to a class with their own notions, ideas, mental models, and beliefs about the world that surrounds them (von Glaserfeld, 1995). Researchers whose work is founded on the tenets of constructivism posit that learning is an active process where the learner constructs or revises their own mental model (Borun, et al., 1993; Murphy & Alexander, 2006; Driver, 1989; Sinatra & Pintrich, 2003). When there is a difference between the mental model a student uses (intuition), and the scientifically accepted explanation, an alternative conception (or misconception) arises (Driver, 1989; Gonen, 2008).

Education research, whether in mathematics or the sciences, attempts to document some common conceptions that students have coming into the classroom (Bar, et al., 1997; Borun, et al., 1993; Clement, et al., 1989; Driver, 1989; Dostal, 2005; Galili & Bar, 1997; Galili, 2001; Graham & Berry, 1993; Watts, 1982). By knowing what students are thinking pre-instruction, educators are in a better position to create the conditions needed to modify a student's conception (if needed) (Dykstra, 2004). That is, educators are in a position to directly address students' misconceptions by creating opportunities for cognitive conflict between ideas (the correct notion in conflict with the student's misconception) and build from that point. Directly addressing student misconceptions is a prominent method used to promote conceptual change in students (Driver, 1989; Greeno, Collins & Resnick, 1996; Murphy & Alexander, 2006; Sinatra & Pintrich, 2003).

Borun, et al.'s (1993) research has shown that people are not able to understand what gravity is until they have seen what it is not. Borun, et al. developed hands-on exhibits that directly confronted persistent misconceptions students have about gravity being associated with air, rotation, and magnetism. In general, the conceptual change approach has its origins in Piaget's theory of equilibration (Kang, Scharmann, Noh, & Koh, 2005) where it is assumed that the need to reduce conflict is human nature and can be key to learning.

By instructing students about the laws of physics, teachers are attempting to fit potentially new knowledge or ideas into the students' existing cognitive structure, or framework. Conceptual modification can take place when students are faced with an example that does not fit into their existing framework. Duit and Treagust (2003) describe an approach to conceptual change that consists of first making the student's framework explicit and then proposing a situation where the framework does not work. The student should become dissatisfied, and at this point a new framework can be proposed. Students are then motivated to accept and modify their existing frameworks. Duit and Treagust (2003) claim that no study has found complete change in student conceptions. Instead, many students develop dual frameworks that are called on in specific situations.

The use of dual frameworks was also documented by Lave (1988) where individuals' use of arithmetic varied drastically with the situation. Given a "real world" situation, like grocery shopping, many of Lave's research participants were able to perform complex arithmetic problems, but when given a formal mathematics test, with similar tasks, the participants' ability plummeted. These differences are classified as situationally specific cognitive activities, or situational knowledge (Greeno, et al., 1996). Understanding that knowledge is situated is an important pedagogical consideration when planning opportunities for learning.

Mildenhall and Williams (2001) refer to well-documented research to validate when students are likely to use a situation-specific approach; in academic contexts, the school-based framework will be used and in "real-world" situations students will fall back to their intuitive frameworks. The school-based and intuitive frameworks do not have to be coherent with one another. Only when both frameworks are called upon does an actual conflict arise (Tall & Vinner, 1981). Plasticity of the mind can allow for a reworking of an existing framework (accommodation), but not entire replacement since these frameworks could have been built up through years of experiences.

Posner, Strike, Hewson, and Gertzog (1982) suggest four conditions that should be met in order for a proposed framework to take the place of an existing framework: dissatisfaction, intelligibility, plausibility, and fruitfulness.   These four conditions were found in most cases where accommodation was documented (Posner et al., 1982). Dissatisfaction can occur when a student has a collection of unsolved puzzles that cannot be explained by current mental models. For the replacement to be considered intelligible, the proposed framework must be non-contradictory and understood by the student. Posner et al. (1982) stress the importance of analogies and metaphors, arguing they lend meaning and intelligibility to new concepts. Further, Chandler supports using a "series of analogies rather than relying on just one" (M. Chadler, personal communication, November 18, 2012). If a proposed framework is to be plausible, then in addition to understanding the framework, the student must find the proposed framework believable. Finally, the new framework will be fruitful if the student finds the framework helpful in solving other problems.

"To understand a physical phenomenon means to know what causes it, what results from it, how to initiate it, how to influence it or how to avoid it" (Greca & Moreira, 1997, pp. 713). In other words, students need some sort of mental model, or framework, to work from in order to understand a physical phenomenon (Besson, 2010). Glynn and Takahashi (1998) define guidelines for creating what they call an elaborate analogy. Glynn and Takahashi found that students who learned about an elaborate analogy not only had greater understanding of the topic, but also had greater recall when compared to students using a standard method of instruction. Their guidelines for an elaborate analogy are (a) the introduction of a target concept, (b) reminding students of the analog concept, (c) identifying the relevant features of the target and analog, (d) mapping the similarities, (e) indicating where/if the analogy breaks down, and (f) drawing conclusions. Glynn and Takahashi also note that making comparisons between the target concept and analogy is sometimes called mapping (1998). The example used in Glynn and Takahashi's research involved both text and images that mapped features from the analogy, a factory, to the target concept, an animal cell. Glynn and Takahashi's results indicate that the analogy acted as a mediator between students existing knowledge and the newly presented knowledge.

Scientific Views of Gravity

Over a long period of time the views held about gravity in the scientific community have changed (Chandler, 1994; Galili, 2001). The Aristotelian, Earth-centered, world view of gravity focused on the quality of matter (color, shininess, amount of earth, air, water, and fire, etc.). The motion of objects in the Aristotelian world was explained by objects trying to reach their "natural place." The natural place of an object was determined by the "nature" of the object; any object not in its proper place would strive to get to the natural place. From an Aristotelian perspective, an unsupported lump of clay, since it is made of earth, properly falls to the ground; being a mixture of air and earth, a feather, will fall more slowly to the earth. Aristotle surmised that more massive objects fall faster. The views of Aristotle were held in high regard for over 2000 years (Hewitt, 2014).

In the 1600s, the scientific community shifted to a Newtonian world view (Chandler, 1994). Chandler (1994) claims the Newtonians viewed the Aristotelian beliefs as superstitious and unscientific. The Newtonians focused on the quantification of data, corresponding with increased popularity of the scientific method. The Newtonian view of gravity involves accepting the gravitational force as being propagated by no physical connection between the two bodies; called action-at-a-distance. Newton was uneasy with the lack of interaction between objects and wrote about it at the end of the *Principia* as

> A certain most subtle spirit which pervades and lies hid in all gross bodies; by the force and action of which spirit the particles of bodies attract one another at near distances, and cohere if contiguous. (as quoted in Chandler, 1994, pp. 165)

Privately, Newton believed that space was an attribute of God (Chandler, 1994). Thus, the force of gravity was a “non-mechanical, spiritual force demonstrating God’s continuing presence, after creation, to maintain the order of the universe” (Chandler, 1994, p. 165). Newton’s law of gravity has been experimentally validated for hundreds of years. Scientists have sent rockets into space relying on this equation alone, even though the equations of General Relativity have surpassed the scope and applications of Newton’s. In certain situations (around large masses or traveling at fractions of the speed of light), we cannot rely on Newton’s gravitational equation alone. In these situations we must take the principles of General Relativity into account. For example, in order for Earth’s GPS devices to work accurately, we must use General Relativity to correct for the differences in rates for clocks in orbiting satellites as compared to the surface of the Earth (Hartle, 2003).

A world view that relies on General Relativity describes gravitational forces as the curvature of four-dimensional spacetime due to the presence of matter (Hartle, 2003). Note that scientists do not know why matter curves spacetime, or even what gravity is, per se. The fundamental meaning of this statement is complex, having implications for the nature of physical law, and how mathematics is used to describe physics. Nevertheless, the connection between gravity and geometry is established fundamentally by Einstein.

How specifically does matter curve space? Beyond insisting that Einstein's equation just says it does, that question may only be answered by extending general relativity into a more comprehensive theory that includes the other pillar of 20th century science, quantum mechanics. In particular, Hartle (2003) notes scientists are still exploring connections between Quantum Gravity and General Relativity. If we look at how scientists' world views have changed over the years, it is reasonable to consider that the General Relativity world view might not be an end-all description of gravity. Like previous perspectives, the General Relativity world view might also be replaced with something else. String Theory is one popular candidate for future descriptions of our universe, but this theory includes many details that need to be verified with experimentation.

While the notion of spacetime curving might seem too extreme for students and the general public to accept, society has already made progress toward this shift of world views. For example, curved spacetime, and even string theory, have been popularized by the media, including many NOVA programs like *The Elegant Universe* (McMaster et al, 2004). More recently, the popularity of Cosmos: A Spacetime Odyssey (Dolleman, 2014), a follow-up to Carl Sagan's Cosmos: A Personal Voyage (Andorfer, 1980), demonstrates the feasibility of introducing the rubber sheet analogy into the classroom.

The Newtonian view of gravity has not, and should not, be completely abandoned. Many problems in physics can be approached from multiple lenses. For instance, some problems in physics are best solved using kinematics, while others can more easily be solved using energy conservation. A physicist has multiple methods at his or her disposal, and is capable of deciding which way to most effectively approach a problem. I argue that introducing students to the rubber sheet analogy allows for another framework students can use to solve problems.

Instruction that includes the rubber sheet analogy may support the formation of a more complete (global) conceptual framework about gravity, one that incorporates both the Newtonian and General Relativity views. In particular, the Newtonian view provides an operational framework that can be employed when students have to make force calculations. The General Relativity view can allow students to accurately solve problems involving GPS technology.

Research into Student Understanding of Gravity

Students' mental models are very resistant to change (Borun, et al., 1993; Limón, 2001; Mildenhall & Williams, 2001; Murphy & Alexander, 2006). Researchers have shown that student conceptions are entrenched and thus are very hard to change because they are supported by every day experiences (Bar, et al., 1997). For example, Bar, et al. (1997) discuss the notion that children regard forces as acting by touch, since that is their usual exposure. Since students feel that contact is needed between the source of the force and the object that it is acting on, a natural place for research into student understanding has been with regard to action-at-a-distance and gravity (Bar, Zinn & Rubin, 1997; Dostal, 2005; Watts, 1982; Watts & Zylbersztajn, 1981). Research has shown that students will fulfill the need for contact to explain gravity by including the presence of air in their explanations (Bar, et al., 1997; Borun, et al., 1993; Watts, 1982). In the absence of air, the gravitational force cannot be transmitted (Bar, et al., 1997; Bradamante & Viennot, 2007; Watts, 1982). Other studies (Bar, et al., 1997; Voutsina & Ravanis, 2013) have found similar associations between the presence of air and the propagation of the magnetic force as well.

Field is a key concept in physics that generally is not introduced until students encounter electricity and magnetism. Greca and Moreira (1997) found that students who do not form mental models of fields show poor conceptual understanding of fields. The standard method of explaining gravity is by using the action-at-a-distance approach, but using action-at-a-distance is not sufficient for extended work with charge interactions. Instead, the idea of fields is introduced to explain the electric force. For example, the field description centers on the idea that space around an electric charge is distorted due to the presence of the charge. The distortion of space results in a vector field, allowing calculations to be done to find the force on a particle in this field. The gravitational field can be visualized by a rubber sheet dipping down around a mass, but the electric field can dip down, as well as up.

Galili (1995) suggests from his results exploring student conceptions of fields in electromagnetism that students find the field idea confusing. To help alleviate student confusion, Galili suggests letting students know the historical reasons why we use the field model, and to also introduce fields earlier when discussing gravity. The rubber sheet analogy is one way to introduce the field model in the context of gravity.

Martinez (2001) summarizes research studies spanning from 1979 to 1995 that have focused on primary and secondary students and their understanding of force. There have been few studies involving college students (Dostal, 2005). Many researchers have classified the mental models students constructed (Galili, 2001; Mildenhall & Williams, 2001) and found that the models generally fall into three categories: an Aristotelian-like intuition, a Newtonian intuition, or some combination of the two. I argue that educators need to add another view of gravity to physics instruction: a view based on General Relativity.

By having complete separation between the two “physics,” educational experiences may lead to student errors and misunderstandings of “even the reality of everyday life, let alone about science” (Galili, 2001, p. 1081). Even though Newton’s Laws allow for sufficient computations in most realms, I argue that the introduction of fields may help learners modify their concept image to be more consistent with the current views of the scientific community. If our goal, as educators, is to develop scientific literacy in students, this additional perspective of spacetime cannot be left out.

While the instruments used to assess student understanding of gravity are vast, there is little consistency in the tools that researchers have developed (Dostal, 2005; Feeley, 2007; Williamson & Willoughby, 2012). This is unlike the Force Concept Inventory (FCI), which many physics education researchers use to gauge student mastery of forces typically taught in a first semester physics course. To assist in the interpretation of the results of my study, I examined normalized gain scores for students taking the FCI. In general, these gain scores have been found to vary between 20 and 59% (Coletta & Philips, 2005). In addition, Coletta and Philips (2005) found that the range of gain scores could be attributed to a student’s reasoning ability, as measured by the Lawson’s Classroom test of Mathematical Reasoning.

Questions used across many studies designed to assess student understanding of gravity, in particular, have common features. Some researchers allow students to answer freely and use drawings to aid in their description (Baldy, 2007; Borun, et al., 1993;

Dostal, 2005; Williamson & Willoughby, 2012), while others rely strictly on a multiple-choice format (Feeley, 2007). One of the critiques by Martinez (2001) in his review of past studies is the lack of evidence supporting the validity of the various instruments. Instruments have been created for each study, but little has been done to expand and replicate studies to other populations. If a common assessment were developed, then much progress could be made in this domain.

A number of researchers have examined possible correlations between performance on science and mathematics assessments and demographic variables (see Kahle & Meece, 1994). Kahle and Meece conducted a meta-analysis of research on classroom performance by sex. The results of the meta-analysis are inconsistent across the studies referenced. For example, some studies show differences in students' achievement in science (Hyde, Fennema, & Lamon, 1990; Stage, Kreinberg, Eccles, & Becker, 1985), whereas other studies have shown no differences (Conner, Schackman, & Serbin, 1978; Liben & Golbeck, 1980; Linn & Peterson, 1985). Also examined were age-related differences in achievement (Hyde, et al., 1990; Hyde, Fennema, Ryan, Frost, & Hopp, 1990). These results are also inconclusive, and thus warrant further investigation. The inconsistent results motivated me to explore possible differences in student performance based on demographic variables.

An Analogy for Gravity

A new approach for teaching gravity was proposed and studied by Baldy (2007) who recommends using Einstein's Theory of General Relativity to explain gravity. General Relativity relies on a geometrical understanding of the shape of spacetime illustrated with the rubber sheet analogy. By using the rubber sheet analogy, educators can initially ignore the action-at-a-distance instructional method and focus student attention on observing how bodies move in response to deformed spacetime. One instructional goal is for students to realize that the motion of an object is not due to some property of attraction inherent in the body, but rather is a response to the deformation of the space in which the object is traveling.

As noted earlier, research has shown students often think an inappropriate medium (e.g. Air) is needed for gravity (Bar, et al., 1997; Borun, et al., 1993; Palmer, 2001; Watts, 1982). As a result, many students think there is no gravity in space or on the Moon (Watts, 1982). I argue that the rubber sheet analogy will give students an anchor to replace students' requirement for a medium, even if their replacement is the rubber sheet itself. There is always a chance that students may develop a different alternative conception that involves the rubber sheet.

Baldy's (2007) study had a small control group that was presented the topic from the traditional Newtonian view. The control group was justified since there is a wealth of data on student understanding in the traditional Newtonian view. The treatment group (*n*=102) consisted of French 9th graders. Students in the treatment group were presented the topic using the Einsteinian view. Students were asked to provide short answers at various points in the lecture sequence in both cases. Student responses were coded and seven concept sectors became apparent, and line up well with prior work. Baldy's study focused on how conceptual change happened at various points in the presentation.

Baldy's (2007) findings indicate that nearly 40% of the students that were exposed to the Einstein approach reached a consistent overall understanding of gravity, 36% of them still confined the phenomenon of attraction to the vicinity of a celestial body, and another 20% thought gravity only occurred on earth. My study aims to expand on Baldy's results by engaging students in hands-on experiences with a physical representation of the analogy.

CHAPTER THREE: METHODOLOGY

I designed my study to assess the effectiveness of the rubber sheet analogy and its ability to affect student conceptions of gravitational fields. I used the following questions to guide my research:

1. Is the rubber sheet analogy an effective model for teaching gravity?
2. What are student conceptions of gravity pre and post instruction?
3. Do personal characteristics or educational background appear to influence student conceptions of gravity?

The research was quantitative in nature; employing a single group, quasi-experimental design with repeated measures. This study was approved by the IRB for exempt status prior to implementation. Students took a pre-test at the beginning of the course, and then a post-test, which was exactly the same as the pre-test, a few weeks after the end of the unit on gravity. The test contained a variety of Likert-scale style and multiple choice questions. I created some questions based on my review of the literature, including the questions that requested demographic information. Regarding demographic information, I was particularly interested in determining students' prior experience with principles of physics and mathematics, age, sex, primary language, class standing, and major.

In addition to the assessments, students participated in a hands-on lab focused on gravity. I developed the lab with the goal of providing students a hands-on opportunity to explore practical implications for gravity using the rubber sheet analogy. The students also had the opportunity to attend multiple lectures on gravity, which included an introduction to the rubber sheet analogy.

In this chapter, I describe my sample population, study design, including the process for creating the instrument and lab, how the analogy was addressed during classroom lectures, and the statistical analyses I used to examine my data.

Participants

The sample in my study was one of convenience. Participants were students enrolled in a 3-credit Conceptual-based Physics course at a 4-year public university. The course had approximately 115 students enrolled, and 84% of the students participated in the pre-test, while 75% of the students completed the post-test. Historically, this class has a majority of first-semester students who have had no formal coursework in physics. The details of gravity as a scientific and mathematical concept was new to them.

All students had the potential to participate in the same treatment. However, student attendance to lecture was not mandatory and was not officially recorded, while student attendance to lab was monitored. Of the students who participated, approximately 40% were female, and 60% male. The course also had a majority of students in the 18-22 age range (59%), and about a third of the class were freshmen standing, a third were sophomore standing, and a third were junior/senior/graduate standing.

Gravity Lab & Equipment Design

I began by trying to build a rubber sheet prototype that would eventually be replicated for the lab set-ups. My initial attempt gave me a product that worked, but was not easily replicable. Since the physics lab sections consist of eight stations that require complete set-ups, I needed a design that was easy to replicate. A break-through in design came about that uses hula-hoops, checkered spandex, and binder clips to fasten (See Figure 2).

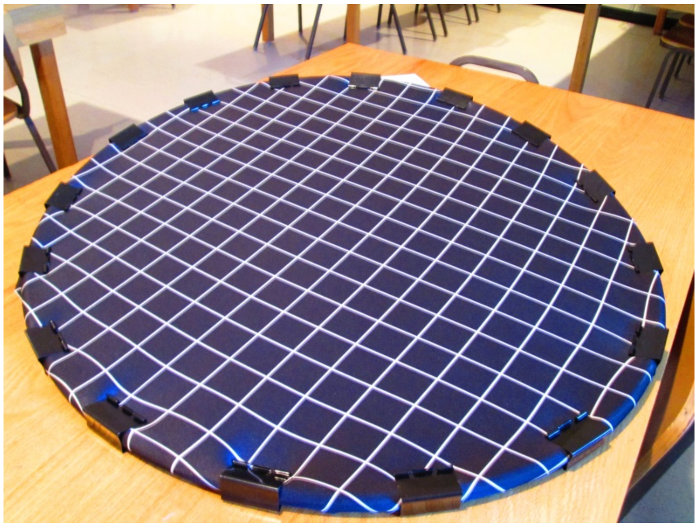


Figure 2. Rubber sheet used in Lab

Since my intent for the lab, in part, was for students to explore the effects on the rubber sheet associated with various masses, my next task was to locate a variety of objects that had similar mass (listed in *Appendix C*) but varying radii, and some masses with same radii but different mass (Figure 3). I developed the lab based on the following instructional goals. Specifically, I wanted student to:

1. Observe how placing various objects with same mass, but varying radii, do not affect the curvature of the sheet.

2. Observe how placing various objects with different mass, but constant radii, change the curvature of the sheet.
3. Observe what happens when two objects are on sheet as the same time. For example, one object is placed in the center and a second ball is either set on the sheet, or given some initial velocity.
4. Reflect and comment on how other parameters might affect the curvature of the sheet (atmosphere, rotation, magnetism).
5. Extrapolate this idea to applications such as gravitational lensing and orbiting.

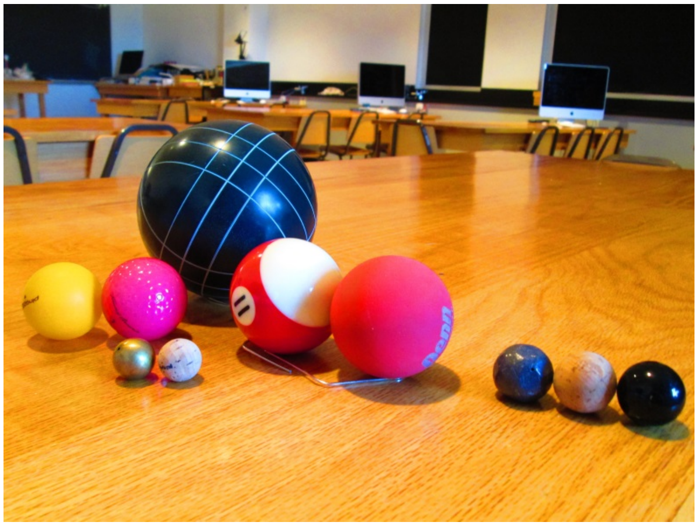


Figure 3. Masses used in Lab

All of these goals are reflected in the final version of the lab (*Appendix A*). I address student and instructor confusion on a few aspects of the write-up in the fifth chapter.

There are five lab sections for the Conceptual-based Physics course involved in my study. As the lead instructor, I taught the first lab section in the week and the other lab sections were taught by upper-level undergraduate physics majors. Lab instructors are expected to review the expectations of each lab prior to the week it occurs and to come to me for clarification when needed. None of the other lab instructors contacted me about the gravity lab prior to its implementation.

Pre/Post-Test Design

After deciding on the objectives for the lab, I began to formulate questions for the instrument I intended to use as my pre/post-test. The initial instrument had items that were adapted from questions found in Williamson and Willoughby (2012) and Dostal (2005). Questions used in both studies allowed students to freely respond. I adapted the most common student responses into multiple-choice options for my instrument. Other questions were adapted from Feeley's (2007) thesis on students' concepts of gravity. I chose the questions so that there were multiple opportunities to assess the same concept.

In the interest of class time, and because I wanted students to take the test seriously, I did not want the test to take too long. My goal was for students to spend no more than 10-15 minutes. Because my initial instrument contained over forty questions, I decided it needed to be pared down. In the interest of creating a test that accurately measures student ideas about gravity, I created a conceptual analysis table (Table 3.1 and Table 3.2) of the test to determine which questions were critical and which could be eliminated (I describe the conceptual analysis in detail in the next section). The pool of questions I chose allows for cross correlation and checking for consistency in answers.

Moreover, I asked experts from education, physics, and diversity programs to review early versions of the instrument and provide feedback to improve overall clarity of expression and ensure accurate presentation of physics concepts. I also tested early versions of the instrument with a variety of individuals ranging in age, gender, and educational background. I made revisions to the instrument after meeting one-on-one with the individuals who reviewed or responded to questions on the instrument.

The final instrument (*Appendix B*) consisted of thirty-seven questions. The first twelve questions are Likert-scale type questions that assess the certainty students have in their responses. Questions thirteen through twenty-eight are multiple choice questions. The final nine questions are demographic questions.

Conceptual Analysis of the Instrument

I conducted a conceptual analysis of the instrument to ensure every item on the instrument was essential (Table 3.1 and Table 3.2) for assessing student understanding of gravity. The instrument included questions and answer choices (distractors) that, when chosen, indicate a students' association with some common concepts that are associated with gravity. I focused on student associations with air, magnetism, and rotation in my study, which are titles of three columns in the table. The titles of the other columns represent other ideas that students may associate with gravity.

The table cell is grayed out if a student might use a concept in reasoning through the question. If the concept is directly used in the statement of the question, then the corresponding table cell also has a Q in it. If the concept was directly addressed in lab, then the cell associated with the appropriate question number has a lined fill. The conceptual analysis table also shows where answer choices directly relate to the concept heading. For example, if I want to see how many students use the concept of rotation to answer a question, then I can look at their responses to either Question 26 or 28. Notice that Question 26 also has several other distractors as answer choices.

Table 3.1

Conceptual Analysis of the Instrument- A

Quest.	Mass	Radius	Falling Objects	Rotation	Atmosp.	Magnetism
▨ 1					Q	
▨ 2				Q		
▨ 3		Q				
▨ 4	Q	■				
5						
6			Q			
▨ 7					■	
▨ 8	■	■	■		■	■
9	■		■			
10			Q			
11						
12	■	■			■	
13	■	■	Q		■	
▨ 14					Ans- D, E	
15	■	■		■	■	■
16						
17	■	■		■	■	■
18	■	■	Q		■	
19	■	■	Q		■	
20			■			
▨ 21				■		Ans- C
▨ 22	Q	■				
▨ 23	Q	Q				
24			■			
▨ 25						Q
▨ 26				Ans- B	Ans- A	Ans- D
▨ 27	■	■			Ans- C	
▨ 28				Ans- A, B		

Table 3.2

Conceptual Analysis of the Instrument- B

Quest.	Distance from Sun	Distance from Surface	Weight vs. Mass	Orbiting	Density
▨ 1					
▨ 2					
▨ 3		■			
▨ 4					■
5	Q				
6					
▨ 7		■	■	■	
▨ 8			■		
9		■	■		
10					
11			Q		
12		■			
13		■			
▨ 14		Q			
15	■		Q		
16			Q		
17	■		■		
18		■			
19		■			
20			■	Q	
▨ 21		Q			
▨ 22					
▨ 23		Q; Ans- C,D			Ans- A, B
24				Q	
▨ 25					
▨ 26					
▨ 27	Ans- B				
▨ 28					

Pre/Post-Test Scoring

I determined each student's total score by contributions from two parts of the test. The multiple choice section (questions 13-28) was scored by awarding one point for each correct response. Thus, students could earn a total of 16 points from the multiple choice section of the test. The second contribution to each student's total score came from the student's responses to questions 1-12. These are the Likert-scale type questions that assess a student's confidence in their answer. Each statement does have a correct answer. To be able to include a student's score on the confidence questions, I gave one point for being certain and having the correct answer, 0.75 points for having the correct answer but only thinking it was right, 0.5 points for choosing option C ("do not know or are uncertain"), 0.25 points for having the wrong answer but only thinking it was right, and 0 points for having the incorrect answer and being certain it was right. Thus, a student could score up to a total of 28 points possible on the test. I used the process described above to calculate a raw score for each student on the pre and post-test.

I also calculated a normalized gain score for paired results using the following standard computation: $100 * \frac{Post-Pre}{100-Pre}$. The normalized gain score measures the relative gain in performance on the test, controlling for students' pre-test scores.

Analogy in the Classroom

The Conceptual-based Physics course that was a part of my study included two 75-minute lectures a week. I gave lectures on the topic of gravity for 1.5 weeks of the course. My lectures included a class discussion of, and calculations with, Newton's Law of Gravity. Collectively, the students and I derived, and calculated, the gravitational acceleration constant, g, for Earth. I also led a class discussion on how the gravitational acceleration constant varies for other bodies, which was followed by students engaging in thought experiments involving the rubber sheet analogy and objects with varying masses. In addition to this, I showed the class videos of the 2-D rubber sheet as well as a simulation of a 3-D rubber sheet. I also assigned homework problems that required the use of a virtual rubber sheet to answer questions. This virtual rubber-sheet was also accessed by students during the gravity lab.

Statistical Analysis

Before each analysis, an *F*-test was performed to confirm that the data sets had similar variances. All tests used a significance level of 0.01. This value was chosen to reduce the potential of a type-1 error (Nuzzo, 2014), even though behavioral sciences often use 0.05 significance levels. Effect sizes are also used in conjunction with the *p*-values to make sure the results are truly meaningful by multiple methods of statistical testing. In order to assess the effectiveness of the analogy, I performed a one-sided *t*-test for independent means on the pre and post-test scores. In addition, I calculated the normalized gain for paired data. Coletta and Philips (2005) completed a study looking at normalized gain scores for students' performance on the Force Concept Inventory. Their analysis only included normalized gain scores between 15 and 80 percent. To be conservative, I examined all but the extreme negative normalized gain scores.

When appropriate, I assessed differences in outcome variables by categorical variables using one-way ANOVA procedures. For example, I compared the pre-test scores for the categories of sex (female and male) to see if their performance was significantly different. I did this again to compare post-test scores.

I used a test for proportion to determine if there was improvement on three specific concepts: rotation, atmosphere, and magnetism. Recall that several questions on the pre/post-test included distractors directly related to one these concepts (atmosphere: 14 and 27; rotation: 28; magnetism: 21 and 25; multiple: 26). By choosing one of these distractors, an association could be made between the students' conception of gravity and the chosen concept. I calculated the test for proportion by determining the standard deviation: $\sigma = \sqrt{\frac{P*(1-P)}{n}}$, where P is the pre-test proportion. The null hypothesis states that the post-test proportion (p) should be equivalent. I calculated the z-score by $z = (p - P)/\sigma$, and then consulted standard z-tables.

To measure the strength of the treatment I also determined effect sizes. In situations where I compared two data sets I calculated the Cohen's *d* effect size and the effect size correlation *r*. When I used ANOVA to test multiple variables, I calculated the eta-squared (η^2) effect size. A small practical significance is associated with effect sizes between 0.20 and 0.50. A moderate practical significance is associated with effect sizes between 0.50 and 0.80. A large practical significance is associated with effect sizes greater than 0.80.

I also calculated Chronbach's alpha test statistic for both the pre- and post-test. This coefficient gives an estimate of internal consistency for the reliability of test scores. The general rule of thumb is that if alpha is between 0.7 and 0.9, the test has good internal consistency. If values are larger than 0.90, the test could be too long or have redundancies.

CHAPTER FOUR: RESULTS

In this chapter, I describe the results of my analysis. I organize the chapter by research question. I discuss these results in Chapter Five. Chronbach's alpha for the pre- and post-test was determined and is 0.86 and 0.85, respectively. These values point to good internal reliability of the instrument.

Is the rubber sheet analogy an effective model for teaching gravity?

I analyzed the average scores from the pre- and post-tests using a *t*-test for independent means. Table 4.1 shows the average score for the pre-test was 51.4%, while the average score for the post-test was 67.8%. There was a significant difference ($p<0.001$) in average scores between the two. Further, Cohen's effect size value ($d=0.89$) suggests a high practical significance. Student performance significantly improved.

Table 4.1

Pre and Post-test Results

	Pre	***Post***
Mean (%)	**51.4**	**67.8**
Variance	**338.8**	**341.4**
n	**97**	**86**
p	**5.1E-09***	
Cohen's *d*	**0.89**	
Effect size *r*	**0.41**	

Table 4.2 contains the statistics associated with the normalized gain for paired students. The average normalized gain was 30.3%. I did not include in the statistical

analysis one student with a normalized gain of -1100%. This student scored a 99.1% on the pre-test and an 89.3% on the post-test.

Table 4.2

Normalized Gain Statistics

Normalized Gain	
Mean (%)	**30.3**
Standard Error	**4.9**
Standard Deviation	**43.9**
Minimum	**-161.1**
Maximum	**93.2**
n	**81**

A histogram (Figure 4) of the normalized gain scores indicates there were only a small number of students with negative gain scores, and that the positive scores are centered around 40. Also notice the histogram shows over half the class had gains at 40% or above.

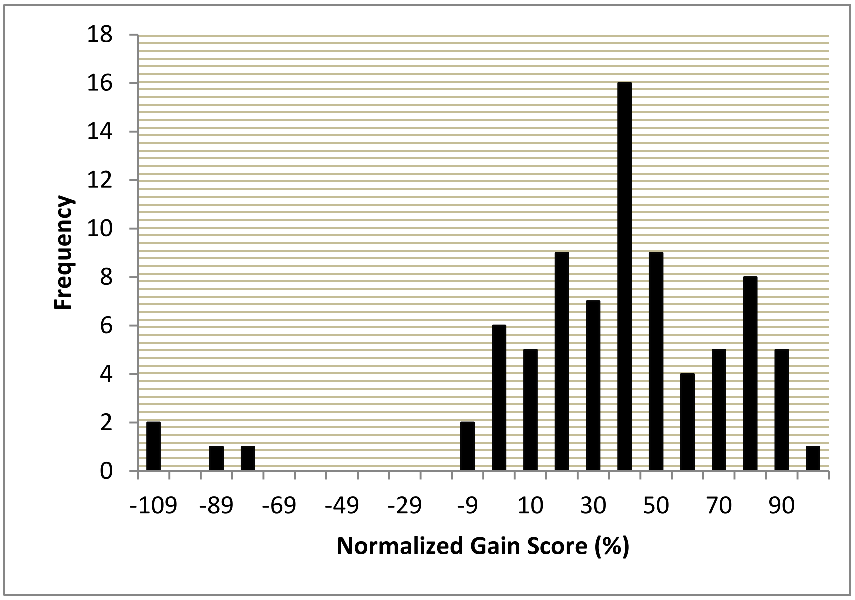


Figure 4. Normalized Gain Histogram

In order to gauge the effectiveness of the gravity lab, I grouped questions from the test that were directly addressed and then I used a *t*-test to compare gains (*p*=0.244) and found no significant difference. I defined the normalized gains associated with questions on the pre/post-test directly related to the lab “very gains” since these are “very related” to tasks completed during the gravity lab. Recall, the conceptual analysis of the instrument (Table 3.1 and Table 3.2) shows which questions are directly related to the lab by the lined texture. The effect size (*d*=0.04) indicates students were able to answer these questions just as well as questions that weren’t directly addressed in the lab (see Table 4.3).

Table 4.3

Gains on questions “very” related to the Lab

	Very gain	***Overall gain***
Mean (%)	**28.4**	**30.3**
Variance	**2455.2**	**1925.9**
n	**81**	**81**
p	**0.244**	
Cohen's *d*	**0.04**	
Effect size correlation *r*	**0.02**	

In an effort to identify whether there was a bias in instruction from my lab to other lab sections, I used a *t*-test to compare post-test scores (see Table 4.4). The scores and variances were not significantly different (*p*=0.081, *d*=0.37). I also used a *t*-test to compare post-test scores from students who completed my lab versus those who did not attend lab at all (see Table 4.5). Again, the scores and variances were not significantly different (*p*=0.211, *d*=0.33). Note that the sample size is rather small for these two analyses. The Cohen’s *d* effect size for these two results are comparable in size and indicate small practical significance.

Table 4.4

Post-test Scores for My Lab/Other Lab Comparison

	My Lab	**Other Lab**
Mean (%)	73.1	66.2
Variance	393.2	323.1
n	19	58
p	0.081	
Cohen's *d*	0.37	
Effect size correlation *r*	0.18	

Table 4.5

Post-test Scores for My Lab/No Lab Comparison

	My Lab	**No Lab**
Mean (%)	73.1	66.7
Variance	393.2	354.4
n	19	9
p	0.211	
Cohen's *d*	0.33	
Effect size correlation *r*	0.16	

I did not include an analysis of individual student lab scores because of the lack in variation in grading across various lab sections. This is due to the fact that a large portion of the lab grade is tied to student attendance.

What are student conceptions of gravity pre and post instruction?

I chose to focus on three concepts from the literature that students often associate with gravity: magnetism, atmosphere, and rotation. I paired the data for pre/post-scores and ran a test for proportion. In particular, I first compared the percentage of students who chose the targeted distractor(s) on the pre-test with the percentage of students who chose the distractor(s) on the post-test. Student performance regarding distractor selection improved significantly on four of the six questions (see Table 4.6). I then compared the percentage of students who chose the correct answer on the pre-test with the percentage of students who chose the correct answer on the post-test. Student performance also improved significantly regarding the proportion of the class that chose the correct answer (see Table 4.7). These results show overall improvement in student performance on the related items (atmosphere: 14 and 27; rotation: 28; magnetism: 21 and 25; multiple: 26). The total percentage between the distractor choice and the correct choice do not necessarily add to 100%. This is due to the fact that there are other possible answer choices that I did not account for in my analysis.

Table 4.6
Distractor Test for Proportion Results

n = 82	**% that Chose Distractor(s)**				
Concept	**PRE**	**POST**	**SD**	**p**	**Effect size *d***
ATMOS.					
Q14	54.9	28	5.5	**< 0.0001***	-0.63
Q27	12.2	2.4	3.6	**0.0034***	-0.95
ROTATION					
Q28	54.9	43.9	5.5	0.0228	-0.24
MAGNETISM					
Q21	9.8	4.9	3.3	0.0681	-0.41
Q25	42.7	26.8	5.5	**0.0018***	-0.39
MULTI					
Q26	82.9	69.5	4.2	**0.0006***	-0.42

Table 4.7
Correct Choice Test for Proportion Results

n = 82	**% of Class Correct**				
Concept	**PRE**	**POST**	**SD**	**p**	**Effect size *d***
ATMOS.					
Q14	20.7	54.9	4.5	**< 0.0001***	0.85
Q27	72.0	82.9	5.0	0.0143	0.35
ROTATION					
Q28	39.0	52.4	5.4	**0.0064***	0.30
MAGNETISM					
Q21	76.8	70.7	4.7	0.0951	-0.17
Q25	50.0	68.3	5.5	**0.0005***	0.42
MULTI					
Q26	17.1	30.5	4.2	**0.0006***	0.42

Students showed significant improvement in both of the atmosphere related questions. Question 14 (Figure 5) asked students to think about how the gravitational force changes as you move up from the Earth’s surface. There were two distractors (answers D and E) that linked gravitational strength to the presence of atmosphere. On the pre-test, 54.9% of the class chose one of these distractors, and on the post-test 28% of the class chose an atmosphere distractor ($p<0.0001$, $d=-0.63$). Class performance on this question went from 20.7% of the students getting this question correct on the pre-test, to 54.9% of the students answering correctly on the post-test ($p<0.0001$, $d=0.85$). The effect sizes suggest moderate to high practical significance related to student performance on Question 14.

14) As you move up and away from the Earth's surface, what happens to the Earth's gravitational force on you?

A. The gravitational force on you decreases, but never reaches zero.
B. The gravitational force on you increases.
C. The gravitational force on you stays the same.
D. The gravitational force on you decreases until you leave the Earth's atmosphere, where it then reaches zero.
E. The gravitational force on you stays the same until you leave the Earth's atmosphere, where it then reaches zero.

Figure 5. Pre/Post-Test Question 14

The second item that involved the relationship between gravity and atmosphere asks students to compare the strength of gravity at the surface of Pluto to the surface of Earth (see Question 27 in Figure 6). There was one distractor (answer C) that linked gravitational strength to the presence of atmosphere. On the pre-test 12.2% of the class chose this distractor, and on the post-test 2.4% of the class chose the atmosphere distractor (p= 0.0034, d=-0.95). The effect size suggests high practical significance for students choosing the atmosphere distractor on Question 27. Class performance on Question 27 went from 72% of the students getting this question correct on the pre-test to 82.9% of the students answering correctly on the post-test (p=0.0143, d=0.35). The effect size shows low practical significance regarding the change in the percentage of students getting the problem correct.

27) How would you compare the strength of gravity at the surface of Pluto with the strength of gravity at the surface of the Earth?

A. Weaker because Pluto has less mass.
B. Weaker because Pluto is further from Sun.
C. Weaker because Pluto has less atmosphere.
D. They are the same.
E. Greater

Figure 6. Pre/Post-Test Question 27

28) The following planets are viewed from above, with more arrows representing a faster rotation. All planets have the same mass and radius. Rank, *from greatest to least*, the strength of gravity on each planet.

A. C=D, B, A because faster rotation creates more gravity.
B. A, B, C=D because less rotation creates more gravity.
C. All have the same gravity.
D. There is not enough information to answer the question.

Figure 7. Pre/Post-Test Question 28

Question 28 asked students to rank the strength of gravity of various planets depending on their indicated rotation (see Figure 7). There were two distractors (answers A and B) that linked gravitational strength to rotation. On the pre-test 54.9% of the class chose one of these distractors, and on the post-test 43.9% of the class chose a rotation distractor (*p*=0.0228, *d*=-0.24). Students did not show significant improvement on this concept. Class performance went from 39% of the students getting this question correct on the pre-test to 52.4% of the students answering correctly on the post-test (*p*=0.0064, *d*=0.30). Even though the *p*-value indicates a significant improvement, the effect size shows low practical significance regarding student performance on Question 28.

Two questions (Figures 8 and 9) had answer choice distractors related to magnetism. Question 21 asked students which person experienced the stronger force of gravity. The image shows a person at the equator, a person at each of the poles, and a person at mid-latitude. There was one distractor (answer C) that indicated the student associated the gravitational strength to magnetism. On the pre-test 9.8% of the class chose this distractor, and on the post-test 4.9% of the class chose the magnetism distractor (*p*=0.0681, *d*= -0.41). While this result is not significant, it is interesting to note for this particular problem 76.8% of the class got this problem correct on the pre-test and only 70.7% of the class answered correctly on the post-test (*p*=0.0951, *d*=0.17). The calculated effect size is 0.17.

21) Which person standing on the surface of the Earth experiences a *stronger force* of gravity?
A. Person C, since they are on the equator.
B. Person A and D, since they are at the poles.
C. Person A since they are at the North magnetic pole.
D. Gravity is the same everywhere on the planet.

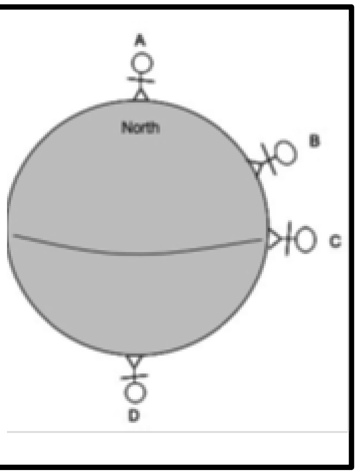
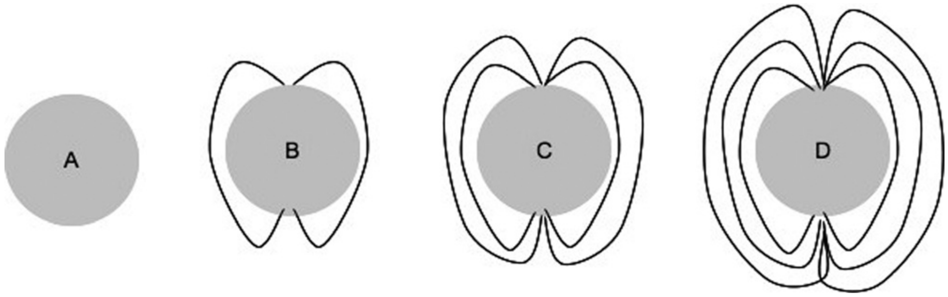


Figure 8. Pre/Post-Test Question 21

Question 25, the other magnetism question, asks students to rank the strength of gravity on each of the planets (see Figure 9). The image accompanying the item shows four planets, all of the same mass and radius, and various black loops representing magnetic fields, where more loops represent stronger magnetic fields. There were two distractors (answers A and B) that linked gravitational strength to magnetism. On the pre-test, 42.7% of the class chose these distractors, and on the post-test 26.8% of the class chose the magnetism distractors (*p*=0.0018, *d*=-0.39). Class performance on Question 25 went from 50% of the students getting this question correct on the pre-test to 68.3% of the students answering correctly on the post-test (*p*=0.0005, *d*=0.42). The effect sizes show low practical significance, but the *p*-values indicate significant results. While this seems like a large improvement, there is still some confusion regarding gravity and magnetism.



25) Rank the strength of gravity (greatest to least) on each of the following planets (if any), where more black loops represent stronger magnetic fields. All planets have the same mass and radius.

A. D, C, B, A because stronger magnetic fields make stronger gravity.
B. A, B, C, D because the magnetic fields cancel out a planet's gravity.
C. All have equal ranking because gravity and magnetism have no relationship.
D. There is not enough information to answer the question.

Figure 9. Pre/Post-Test Question 25

The final question I analyzed was Question 26 (see Figure 10). Question 26 asks students to identify if there are any other forces, besides gravity, that hold us to the Earth's surface. Answer choices are "atmospheric forces," "rotational dynamics," "friction," "magnetic forces," "more than one of the previous choices," or "nothing else." Choosing anything but the correct choice of "nothing else" indicated that the student still has some misconception about gravity. On the pre-test 82.9% of the class chose one of these distractors, and on the post-test 69.5% of the class chose a distractor (*p*=0.0006, *d*=-0.42). Class performance on this question went from 17.1% of the students getting this question correct on the pre-test to 30.5% of the students answering correctly on the post-test (*p*=0.0006, *d*=0.42). Again the effect sizes show low practical significance, but the *p*-values are significant. Thus, many students still incorrectly associated some unrelated idea to gravity.

26) Besides gravity, are there any other forces that hold us to the Earth's surface?
A. Atmospheric forces.
B. Rotational dynamics.
C. Friction.
D. Magnetic Forces.
E. More than one of the above factors.
F. Nothing else.

Figure 10. Pre/Post-Test Question 26

Do personal characteristics or educational background appear to influence these conceptions?

There are eight demographic questions on the pre and post-test. These questions are typical to educational research and were included to make it possible to explore potential performance differences between various groups. I performed a *t*-test for independent means and two statistically significant results emerged for student performance based on age and sex (see Table 4.8).

Table 4.8

Results for Questions 29 & 30

Question/Response		**Pre- Test**			**Post-Test**	
Age	***n***	**M**	**SD**	***n***	**M**	**SD**
18-22	56	47	14.7	54	63.8	18.3
23 & over	39	57.4	20.8	31	74.6	17.3
p		**0.0025***			**0.0044***	
Cohen's *d*		**0.58**			**0.61**	
Effect size correlation *r*		**0.28**			**0.29**	
Sex	***n***	**M**	**SD**	***n***	**M**	**SD**
Female	38	42.5	13.8	34	59.7	18.7
Male	56	57.3	18.5	51	73.1	16.6
p		**< 0.0001***			**0.0004***	
Cohen's *d*		**0.91**			**0.76**	
Effect size correlation *r*		**0.41**			**0.35**	

Ages were grouped into 18-22 and 23 and over. These two groups performed consistently different from each other pre (*p*=0.0025, *d*=0.58) and post instruction/lab (*p*=0.0044, *d*=0.61), but had similar gains in average percentage points. The effect sizes indicate moderate practical significance, while the *p*-values indicate significant differences in both the pre-test and post-test scores. The older age group performed significantly better.

The second significant result came from comparing male to female gains. Females started out far behind the males (*p*<0.0001, *d*=0.91), had similar gains in average percentage points, but the females' post-test scores (*p*=0.0004, *d*=0.76) were slightly above the pre-test scores for males. The effect sizes indicate high practical significance, and the *p*-values indicate significant differences in both the pre-test and post-test scores. The males consistently performed significantly better.

In my exploration of native and non-native English speakers, the statistical analyses did not reveal a statistically significant difference between the groups. On the pre-test, non-native English speakers (50.5%) performed similarly to native English speakers (52.3%) (p=0.66, d=0.09). On the post-test, non-native English speakers (60.9%) showed much less gain in scores when compared to the native English speakers (70.2%) (p=0.06, d=0.24). The average percentage-point gain for non-native English speakers was 10.4%, while native English speakers had an average percentage-point gain of 19.7%.

I conducted an ANOVA for all other demographic questions and there were no other significant results (see *Appendix D*). For example, students with prior physics experience and high levels of mathematics did not perform any better than students with no prior experience or low levels of mathematics achievement.

CHAPTER FIVE: DISCUSSION

In this chapter, I discuss the relevant results. The discussion is organized by my research questions. In addition, I address instructional implications and propose potential directions for future research.

Interpreting the Results

Based on the pre/post-test results, students' understanding of gravitational fields had an average gain of 30.3%. Related literature indicates this is a moderate gain for student performance (Coletta & Philips, 2005). If I take a more conservative cut of my data, and limit the normalized gain scores to only positive values, the average normalized gain score for the class increases to 41.9%, a value closely aligned with the peak on the histogram in Figure 4. This is not unusual to do; in the study completed by Coletta and Phillips (2005), normalized gain scores on the FCI were computed, and analysis was limited to only scores between 15 and 80 percent.

Interpretation of the normalized gain hinges significantly on the students reasoning ability. Coletta and Philips (2005) claim that only one-third of college students are able to reason abstractly and scientifically. These formal reasoning skills are required for the study of physics. Kang, et al., (2005) found that a student's logical thinking ability was a significant predictor of achievement level on their study's conception test. It is possible that if a student is stuck in a transitional stage of reasoning, then there is little hope the student can fully appreciate the intricacies of a topic discussed in physics class. Limón (2001) also argues that radical conceptual change might not be possible at all developmental stages. If this is true for the participants of my study, then interpretation of my results and the assessment of the overall approach of using the rubber sheet analogy to teach gravity is complicated. Since I did not test the reasoning ability of the class, I cannot claim to know my students' reasoning ability. The histogram (Figure 4) roughly shows 1/3 of the class achieving high gain scores. This is consistent with 40% of Baldy's (2007) students achieving an overall global understanding of gravity where students were able to extrapolate principles of gravity beyond a Earth-centered context.

One way I attempted to gauge the effectiveness of the gravity lab was to compare student gain scores for questions highly correlated with the concepts addressed in the lab versus overall gain scores. The effect size was 0.04, indicating low practical significance. These results could imply that students were capable of transferring their knowledge to more situations than just those focused on in lab.

Another way I attempted to gauge the effectiveness of the gravity lab was to compare post-test scores for various lab students. For the few students who did not complete the lab, results indicate that completion of the lab did not make a difference in student performance on the post-test. The average score for "No Lab" (66.7%, *n*=9) versus "Other Lab" (66.2%, *n*=58) were not significantly different. I am not sure who these nine students are who did not attend lab. There are many scenarios that could explain the lack of difference. Without additional information about these students, it is difficult to explain. Of course, this could be an indication the lab was not effective.

Students also improved in their associations with gravity and other factors often deemed influential such as atmosphere, rotation, and magnetism. Table 4.6 shows that the proportions of students who answered correctly was significant for at least one question directly related to each concept. Fewer students also chose the specific distractor at a significant level.

The results from the atmosphere section had high effect sizes. For Question 14, more of the class chose the correct choice at a significant effect level (*d*=0.85), but 45.1% of the class was still choosing an incorrect option. Question 27 had a significant decrease in the number of students choosing the distractor (*d*=0.95), but the percent of the class who answered correctly did not increase significantly. These two questions have very different contexts, which could explain the differences in results. The context of Question 14 is Earth, while the context of Question 27 is Pluto.

Rotation was a concept that did not show significant improvement overall. A large portion of the class still chose the rotation distractor (43.9%) on the post-test. Thus, a significant portion of the students who took the post-test still associated rotation with gravity.

The concept of magnetism also had an interesting result. Question 21 did not have significant results for either analysis. The class actually scored worse on the post-test when looking at the percent correct (76.7% to 70.7%). This could be because students might perceive answers A and B to be associated with rotation, which is a concept close to half the class still associates with gravity. Regardless of the explanation, student conceptions appear to be entrenched and hard to change.

It is interesting to note that even though the class did improve on most conceptions, on Question 26 close to 70% of the class still associated some other factor with gravity. These associations seem to persist in the class even after explicit lecturing, videos, and a hands-on lab experience. This is similar to the results found by Baldy (2007) in her study using the Einstein approach. Baldy found that approximately 56% of her sample still thought of gravity on a local scale (unique to Earth or other celestial bodies).

It is unfortunate that female students were so far behind their male counterparts. They start out about 20 percentage points behind, have similar gains in percentage points, but barely surpass the mean of male pre-test scores. More information is needed to better understand the factors influencing the gap in performance between sexes.

Also, the comparison of the two age groups shows significant differences between. This result demonstrates older students generally had a better scientific understanding of gravity, which is consistent with other studies (Bar, et al., 1997; Borun, et al., 1993; Bradamante & Viennot, 2007; Palmer, 2001). My findings could also be related to the formal reasoning ability of each group, and could explain why the older students have higher scores.

While the results from comparing non-native versus native English speakers were not statistically significant, the percentage point gain between the pre- and post-test for non-native English speakers was about half that of the English speakers (10.4% versus 19.7%). This is potentially relevant for most university instructors because of the large number of refugee and international students attending American universities. There has been a 40% increase in international students attending schools in the American university system over the past decade (DeSilver, 2013). International students account for about 4% of the student population (DeSilver). More research is needed to explain why the non-native English speakers who participated in my study did not have comparable average percentage point gains when compared to native English speakers.

Limitations

One limitation to this study is the lack of communication with the lab instructors. For future iterations, I recommend instructors take the necessary steps to make sure all lab instructors understand every aspect of the lab write-up. One possibility would be to require lab instructors to attend a short meeting the week prior to lab. During this meeting, it may also be helpful to include a discussion of the pre-/post-test.

Another limitation of this study pertains to questions on the lab write-up. For example, students communicated confusion on question B3, which asked them about the physical set-up required to test relativity by measuring the apparent shift in location of a star as its light passes near the Sun. No one in my lab section was able to answer this question without help. I took some time in the middle of lab to talk about this and draw a diagram of the physical set-up. When I went to visit another lab to see how they were doing, there was confusion on this part too. The confusion was not set straight by the lab instructor because he did not know what the question was asking either. He had interpreted it to be somehow related to a Doppler shift.

Further development of the lab should challenge some of the lasting beliefs students have about atmosphere, rotation, and magnetism. I propose adding a hands-on component directly related to these misconceptions. For example, students could use magnetic balls, or rotate objects. This would likely make the lab more useful in changing students' common connections to atmosphere, rotation, and magnetism and could be tested by noting test score differences in future iterations.

There were limitations to my study associated with the pre/post-test. I did not establish external validity and reliability of the instrument. However, prior to initiating further tests for validity and reliability, the instrument could also benefit from some additional editing. Most importantly is the addition of specific questions that require students to use the rubber sheet analogy, which I have not been able to find in the literature and thus will need to be developed. All studies that I have found solely rely on testing the Newtonian aspect of gravity. Revision of a few questions is also needed. For instance, Question 3 should more explicitly state that the observer is located on the surface (like I had intended), and not at some far off point (like some might interpret).

Interpretation of the normalized gain scores was limited because I did not include a measure of formal reasoning ability. Knowing the percentage of students that have reached a formal reasoning stage would inform expectations for possible gains. This information could help indicate the effectiveness (or lack of effectiveness) of the rubber sheet analogy.

Instructional Implications

There is evidence that the rubber sheet analogy with accompanying lab activities is worth continued pursuit. The gains in performance, coupled with reductions in misconceptions, suggest the rubber sheet analogy is worth considering for future instruction. The analogy could provide the scaffold upon which students need to attach their existing ideas, and therefore it could be helpful in laying a good conceptual foundation.

Additionally, using the rubber sheet analogy in an astronomy class would allow for many more applications to be discussed. Educators could use the rubber sheet to talk about, and demonstrate, orbits, gravitational lensing, planetary formation, and black holes. Also, testing this method in various contexts would add to the sparse literature on this topic and address the external validity of the instrument.

I suspect that more explicit attention needs to be given to the three commonly associated concepts (atmosphere, rotation, and magnetism) since approximately 70% of the class still associated some other factor besides mass and separation distance. It is important for educators to know the common conceptions in their classroom, especially since the above list is not exhaustive. Students in my class watched videos, participated in lecture, and completed a hands-on lab, and still the misconceptions persist.

Students used a virtual rubber sheet in a portion of the gravity lab and in their online homework. This applet simulates what happens to the rubber sheet as one changes the radius for the same given mass. It has a slider bar that allows the user to smoothly change through a large span of masses. In the future, I plan on more extensive use and tracking of the virtual rubber sheet applet through student homework. The online homework system that uses this applet will easily allow for further analysis of student responses. This is especially important since these questions are directly related to students’ use of the analogy in solving problems.

Future Research

Research into student understanding of gravity may be expanded by including opportunities for student interaction with the rubber sheet analogy and associated lab across more sections of physics and astronomy courses. Such follow up studies are necessary to determine if age and sex are independent from understanding gravity concepts. A more in-depth analysis that compares the performances of various demographic groups (young women/men to old women/men) could be insightful. Would we still see an achievement gap from this type of higher-order analysis?

With refinement, the instrument I developed for my study could serve as a standard assessment of gravity, similar to how the FCI is the standard for student understanding of force. With a revised instrument, I intend to expand this work to other courses.

Since there is little research into using the rubber sheet analogy, expanding my study to include different ages, as well as courses, could be useful to the education community. It could add more documentation about student misconceptions associated with gravity. Additional research regarding the use of the rubber sheet analogy could also demonstrate the effectiveness of hands-on lab activities and their ability to affect change in student understanding.

The intent of my study was to assess the potential effectiveness of the rubber sheet analogy and its ability to affect student understanding of gravitational fields. I developed instructional materials, including a hands-on lab, to engage university students in thinking about gravity using the rubber sheet analogy. The results showed promise in using this approach in the classroom and should be further explored.

REFERENCES

References marked with an asterisk indicate studies included in the meta-analysis.

Andorfer, G. (producer). (1980). Cosmos [Documentary]. U.S.A.: KCET, Carl Sagan Productions.

Baldy, E. (2007). A new educational perspective for teaching gravity. *International Journal of Science Education*, *29*(14), 1767-1788. doi: 10.1080/09500690601083367

Bar, V., Zinn, B., & Rubin, E. (1997). Children's ideas about action at a distance. *International Journal of Science Education*, *19*(10), 1137-1157. doi: 10.1080/0950069970191003

Besson, U. (2010). Some features of causal reasoning: common sense and physics teaching. *Research in Science & Technological Education*, *22*(1), 113-124. doi: 10.1080/0263514042000187575

Borun, M., Massey, C., & Lutter, T. (1993). Naive knowledge and the design of science museum exhibits. *Curator: Quarterly Publication of the American Museum of Natural History*, *36*, 201-219.

Bradamante, F., & Viennot, L. (2007). Mapping gravitational and magnetic fields with children 9-11: Relevance, difficulties and prospects. *International Journal of Science Education*, *29*(3), 349-372. doi: 10.1080/09500690600718245

Chandler, M. (1994). Philosophy of gravity: Intuitions of four-dimensional curved spacetime. *Science & Education*, *3*, 155-176.

Clement, J., Brown, D., & Zietsman, A. (1989). Not all preconceptions are misconceptions: finding 'anchoring conceptions' for grounding instruction on students' intuitions. *International Journal of Science Education*, *11*(5), 554-565.

Coletta, V. P., & Philips, J. A. (2005). Interpreting FCI scores: normalized gain, preinstruction scores, and scientific reasoning ability. *American Journal of Physics, 73*(12), 1172-82. doi:10.1119/1.2117109

* Conner, J. M., Schackman, M., & Serbin, L.A. (1978). Sex-related differences in response to practice on visual-spatial test and generalization to a related test. *Child Development*, *49*, 24-29.

DeSilver, Drew (2013, November 12). Record number of international students studying in U.S.. *Pew Research Center*. Retrieved from http://www.pewreaserchcenter.org

Dolleman, E. (co-producer). (2014). *Cosmos: The Space Odyssey* [Documentary]. U.S.A.: Cosmos Studios, Fuzzy Door Productions, National Geographic Channel, Six Point Harness.

Dostal, J. A. (2005). *Student concepts of gravity*. (Master's Thesis). Retrieved from http://www.physicseducation.net/members/Dostal_Thesis.pdf

Driver, R. (1989). Students' conceptions and the learning of science. *International Journal of Science Education, 11*(5), 481-490.

Dykstra, D. (2004). "How We Think About and Prepare to Teach Physics." *Proceedings of the 2004 Physics Education Research Conference*, *790*209-212. http://dx.doi.org/10.1063/1.2084738

Duit, R., & Treagust, D. F. (2003). Conceptual change: A powerful framework for improving science teaching and learning. *International Journal of Science Education*, *25*(6), 671-688. doi: 10.1080/09500690305016

Feeley, R. E. (2007). *Identifying student conceptions of gravity*. (Master's thesis).

Galili, I. (1995). Mechanics background influences students' conceptions in electromagnetism. *International Journal of Science Education*, *17*(3), 371-387. doi: 10.1080/0950069950170308

Galili, I. (2001). Weight versus gravitational force: Historical and educational perspectives. *International Journal of Science Education*, *23*(10), 1073-1093. doi: 10.1080/09500690110038585

Galili, I., & Bar, V. (1997). Children's operational knowledge about weight. *International Journal of Science Education, 19*(3), 317-34-. doi: 10.1080/0950069970190305

Glynn, S. & Takahashi, T. (1998). Learning from analogy-enhanced science text. *Journal of Research in Science Teaching, 35*(10), 1129-1149.

Gonen, S. (2008). A study on student teachers' misconceptions and scientifically acceptable conceptions about mass and gravity. *Journal of Science Education and Technology , 17*(1), 70-81. doi: 10.1007/s10956-007-9083-1

Graham, T. & Berry, J. (1993). Students' intuitive understanding of gravity. *International Journal of Mathematical Education in Science and Technology, 24*(3), 473-478. doi: 10.1080/0020739930240319

Greca, I. M. & Moreira, M. A. (1997). The kinds of mental representations- models, propositions and images- used by college students regarding the concept of field. *International Journal of Science Education, 19*(6), 711-724. doi: 10.1080/0950069970190607

Greeno, J.G., Collins, A.M. & Resnick, L.B. (1996). Cognition and learning. In: Berliner, D.C. & Calfee, R.C.. Handbook of Educational Psychology. New York: Prentice Hall International, 15-46.

Hartle, J. (2003). *Gravity: An introduction to einstein's general relativity*. (1st ed.). San Francisco: Addison Wesley.

Hewitt, P. (2014). *Conceptual physics*. (12th ed.). San Francisco: Pearson.

* Hyde, J.S., Fennema, E. & Lamon, S.J. (1990). Gender differences in mathematics performance: A meta-analysis. *PsychologicalBulletin, 107*(2), 139-155.

* Hyde, J.S., Fennema, E., Ryan, M., Frost, L., & Hopp, C. (1990). Gender comparisons of mathematics attitudes and affect: A meta-analysis. *Psychology of Women Quarterly, 14*(3), 299-324.

Kahle, J., & Meece, J. (1994). Research on gender issues in the classroom. In D. Gabel & National Science Teachers Association (Eds.), *Handbook of research on science teaching and learning.* New York: Macmillian.

Kang, S., Scharmann, L., Noh, T., & Koh, H. (2005). The influence of students' cognitive and motivational variables in respect of cognitive conflict and conceptual change. *International Journal of Science Education*, *27*(9), 1037-1058. doi: 10.1080/09500690500038553

Lave, J. (1988). Cognition in practice: Mind, mathematics, and culture in everyday life. Cambridge University Press.

* Liben, L.S., & Golbeck, S.L. (1980). Sex differences in performance on Fiagetian spatial tasks: Differences in competence or performance?, *Child Development*, *51*(2), 594-597.

Limón, M. (2001). On the cognitive conflict as an instructional strategy for conceptual change: A critical appraisal. *Learning and Instruction*, *11*, 357-380.

* Linn, M.C., & Peterson, A.C. (1985). Facts and assumptions about the nature of sex differences. In S.S. Klein (Ed.), *Handbook for achieving sex equity through education* (pp. 53-77). Baltimore: Johns Hopkins University Press.

Martinez, N. M. (2001). Characteristics of the methodology used to describe students' conceptions. *International Journal of Science Education*, *23*(7), 663-690. doi: 10.1080/09500690010006509

McMaster, J., Cort, J., Greene, B., Howes, R., WGBH (Television station: Boston, Mass.), Channel Four (Great Britain), David Hickman Films (Firm), WGBH Video (Firm). (2004). *The elegant universe*. Boston: WGBH Boston Video.

Mildenhall, P. T., & Williams, J. S. (2001). Instability in students' use of intuitive and Newtonian models to predict motion: The critical effect of the parameters involved. *International Journal of Science Education*, *23*(6), 643-660. doi: 10.1080/09500690010006545

Murphy, P. K., & Alexander, P. A. (2006). *Understanding how students learn*. (1st ed.). Thousand Oaks: Corwin Press.

Nuzzo, R. (2014). Statistical errors: p-values, the 'gold standard' of statistical validity, are not as reliable as scientist might assume. *Nature*, *506*, 150-152.

Palmer, D. (2001). Students' alternative conceptions and scientifically acceptable conceptions about gravity. *International Journal of Science Education*, *23*(7), 691-706. doi: 10.1080/09500690010006527

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education*, *66*(2), 211-227. Retrieved from ci.unlv.edu/files/Week6_Posner_et_al_Optional.pdf

Sinatra, G. M., & Pintrich, P. R. (2003). *Intentional conceptual change*. Mahwah, N.J.: L. Erlbaum.

* Stage, E.K., Kreinberg, N., Eccles, J., & Becker, J. (1985) Increasing the participation and achievement of girls and women in mathematics, science, and engineering. In S.S. Klein (ed.), *Handbook for achieving sex equity through education* (pp. 237-268). Baltimore: Johns Hopkins University Press.

Tall, D., & Vinner, S. (1981). Concept image and concept definition in mathematics with particular reference to limits and continuity. *Educational Studies in Mathematics*, *12*(2), 151-169. Retrieved from http://www.jstor.org/3482362

von Glaserfeld, E. (1995). *Radical constructivism: A way of knowing and learning*. London: Taylor & Francis Ltd.

Voutsina, L., & Ravanis, K. (2013). Magnetism and gravity: Mental representations of students 15-17 years old from a historical and teaching perspective. *Journal of Social Science Research*, *1*(3), 49-57.

Watts, D. M. (1982). Gravity- don't take it for granted!. *Physics Education*, *17*, 116-121. Retrieved from http://iopscience.iop.org/0031-9120/17/3/306

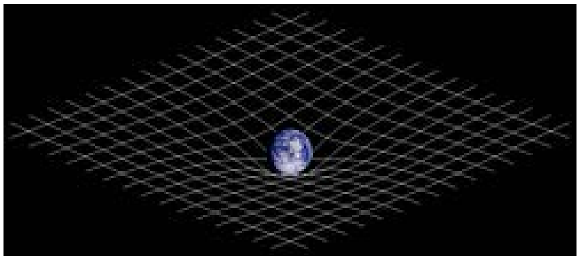
Watts, D. M., & Zylbersztajn, A. (1981). A survey of some children's ideas about force. *Physics Education*, *16*, 360-365. Retrieved from http://iopscience.iop.org/0031-9120/16/6/313

Williamson, K. E., & Willoughby, S. (2012). Student understanding of gravity in introductory college astronomy. *Astronomy Education Review, 11*(1), 105. doi: 10.3847/AER2011025

APPENDIX A

Gravity and Spacetime Lab

Gravity and Spacetime



Humans are not very good at visualizing situations in four dimensions. Instead, a two dimensional analogy can be used when describing the curvature of spacetime. It is called the rubber sheet analogy. Imagine a large stretchy sheet of rubber pulled out to be completely flat.

What will happen when you place different objects onto the sheet? How will multiple objects interact? The curvature of the rubber sheet is representative of how matter distorts the space around it, and objects follow this distorted space. This is what we perceive as a gravitational force, and is distinctly different from Newton's way of describing gravity.

Equipment: A large stretchy sheet, various sized balls (some should vary in radius while holding the mass constant, and vice versa), and internet access.

A. Introduction

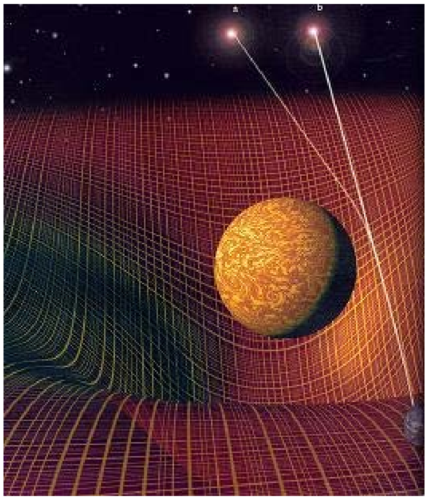
A1. Gather a variety of masses. Place each on the rubber sheet. Describe what happens when:

- The mass increases
- The radius increases
- More than one object is located on the sheet
- Compare effects from two objects of equal mass, but different radii (or same radii, but different mass)

A2. The strength of the gravitational force between two objects can be calculated using Newton's Universal Gravitational force law. It states the force is directly proportional to the mass of the two objects. *How does this relate to your above observations?*

A3. The strength of the gravitational force also depends on how far apart the two objects are from one another. It has an inverse-square relationship. *What does this mean and how does it relate to your above observations?*

B. Verification



In 1916 Einstein proposed three classical tests of general relativity: precession of Mercury's orbit, gravitational redshift, and deflection of light by the Sun. We will focus on the last.

B1. Select a mass and place it in the center of your sheet. This mass will act as your star. Practice rolling a second ball so that it goes around it without a collision. *What happens to the path of the rolling ball as it passes the star?*

B2. Now vary the mass of the star. Try at least three different masses. *What happens to the rolling balls path as the mass of the star increases?*

B3. As the above image shows, this deflection of the rolling mass makes it appear to come from another location. Measuring the apparent shift in location of a background star was done in 1919. *What are the required conditions on Earth for an observer to verify this? Why is this so?*

During the 1919 solar eclipse, experimenters in various locations across the globe made measurements of this deflection. The success of this experiment was published in newspapers around the world and raised Einstein and his theory to rock star status.

B4. The object that bends spacetime does not have to be a star. Any mass distorts the fabric of spacetime and can cause changes to the path of light passing through it. In general, this method is called Gravitational Lensing. It can cause an increase in light from the background object. *Can you name other objects, besides a single star, that can bend light that astronomers might use?*

B5. Watch the short film created by The Max Plank Institute of Astrophysics regarding Gravitational Lensing: http://www.youtube.com/watch?v=yamVbK-J69M. Our rubber sheet analogy is two-dimensional. *What kind of images do we actually get from our four-dimensional observations?*

C. Virtual Rubber Sheet

C1. Go to the virtual lab:
http://media.pearsoncmg.com/aw/aw_0media_astro/if/SWF/Spacetime_Mass_Rad_Orbit.swf

Describe what happens when:

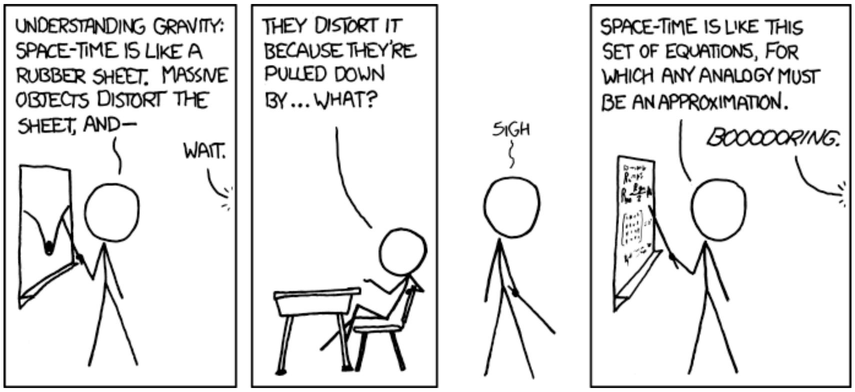
- The mass increases
- The radius increases

Is this consistent with your answers in part A? If not, what is different?

C2. How do each of the following parameters effect the curvature of space, and hence, the strength of gravity on a planet's surface?

- The mass of the planet
- The atmosphere of the planet
- The radius of the planet
- The rotation rate of the planet
- The planets magnetic field

C3. *What would happen to Earth's orbit if the Sun became a black hole?*



Permanent link to this comic: http://xkcd.com/895/

APPENDIX B

Pre/Post-Test for Study

Please bubble in your LAST NAME, FIRST NAME on the provided scantron. Your answers will help inform our later lectures on gravity. You will not be graded on this. Your name will be removed before the data is given to Tiffany Watkins.

Use the following for questions 1 through 12:
If you are *certain* it is **true**, choose **A**.
If *you think* it is **true**, but are not so sure, choose **B**.
If you do not know, or are uncertain, choose **C**.
If *you think* it is **false,** but are not sure, choose **D**.
If you are *certain* it is **false**, choose **E**.

1) A B C D E **A planet's atmosphere affects its gravitational pull.**

2) A B C D E **A planet's rate of rotation affects its gravitational pull.**

3) A B C D E **A planet's radius affects its gravitational pull.**

4) A B C D E **A planet's mass affects its gravitational pull.**

5) A B C D E **A planet's distance from the Sun affects its gravitational pull.**

6) A B C D E **A planet's gravitational pull affects how fast an object falls towards its surface.**

7) A B C D E **There is no gravity in outer space.**

8) A B C D E **There is no gravity on the moon.**

9) A B C D E **In low gravity, some objects may be too light to be affected by the gravitational force.**

10) A B C D E **Gravity affects how fast an object falls.**

11) A B C D E **Gravity** *does not* **affect the weight of an object.**

12) A B C D E **Heavy objects are hard to lift because Earth's gravitational force increases as you lift.**

13) Suppose you were standing on the moon holding an apple. If you were to let go of the apple, in what direction will it move?

A. The apple will float upward from the lunar surface.
B. The apple will float around, staying about the same height and location.
C. The apple will float around, but also move away horizontally.
D. The apple will fall toward the lunar surface.
E. Other/None of the above.

14) As you move up and away from the Earth's surface, what happens to the Earth's gravitational force on you?

A. The gravitational force on you decreases, but never reaches zero.
B. The gravitational force on you increases.
C. The gravitational force on you stays the same.
D. The gravitational force on you decreases until you leave the Earth's atmosphere, where it then reaches zero.
E. The gravitational force on you stays the same until you leave the Earth's atmosphere, where it then reaches zero.

<u>Use the following for questions 15 through 18:</u>
Venus is sometimes called Earth's sister planet, and is the second planet from the Sun. It is nearly the same size and mass, but Venus rotates once on its axis every 243 Earth days, has an atmospheric pressure 90 times that of the Earth, and practically no intrinsic magnetic field.

15) If you could weigh yourself on Venus, using a standard bathroom scale, you would weigh

A. A lot more.
B. A lot less.
C. About the same.
D. Exactly the same.
E. There is not enough information to answer the question.

16) Your mass on Venus would be

A. A lot more.
B. A lot less.
C. About the same.
D. Exactly the same.
E. There is not enough information to answer the question.

17) The gravitational force of Venus is _____ the gravitational force of Earth.

A. Much greater than
B. Much less than
C. About the same as
D. Exactly the same as
E. There is not enough information to answer the question.

18) Suppose you let go of an apple while standing on the surface of Venus. Compared to releasing an identical apple at the same height while standing on the surface of the Earth, the apple on Venus

A. Will hit the ground in much less time (fall a lot faster) than the apple on Earth.
B. Will hit the ground in much greater time (fall a lot slower) than the apple on Earth.
C. Will hit the ground in about the same amount of time (fall about the same way) as the apple on Earth.
D. Will hit the ground in exactly the same amount of time (fall exactly the same way) as the apple on Earth.
E. Will not fall.

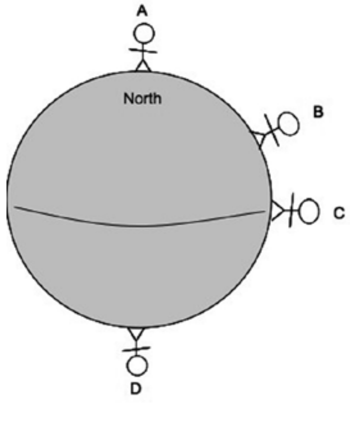
19) It is the year 2156 and people are living on the surface of the moon inside giant domes. These domes are filled with air so that people can live inside the dome without having to wear space suits. Suppose someone is standing inside one of the domes, with an apple in hand. What will happen to the apple if they let go of it?

A. The apple will float upward from the lunar surface.
B. The apple will float around, staying about the same height and location.
C. The apple will float around, but also move away horizontally.
D. The apple will fall toward the lunar surface.
E. Other/None of the above.

20) Why do astronauts appear to float in their spacecraft while orbiting?

A. There is no gravity in space.
B. Gravity is much weaker in space.
C. They are too far away from Earth or any massive body.
D. They are in a constant state of free-fall.
E. The spacecraft's gravity isn't strong enough.

21) Which person standing on the surface of the Earth experiences a *stronger force* of gravity?



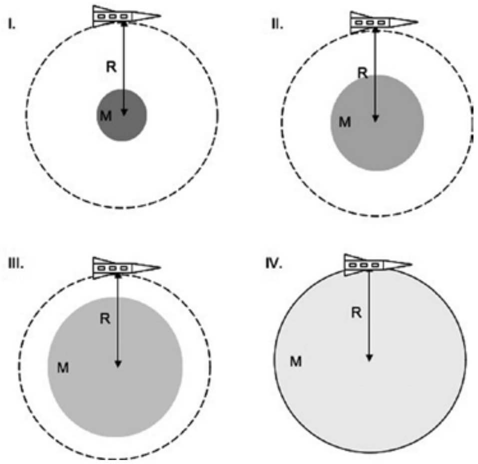
A. Person C, since they are on the equator.
B. Person A and D, since they are at the poles.
C. Person A since they are at the North magnetic pole.
D. Gravity is the same everywhere on the planet.

22) How heavy or light does something have to be to create its own gravitational field?

A. Every object has some gravitational field.
B. Very, very heavy.
C. Planet or moon sized.
D. It needs to at least have a weight I can detect.

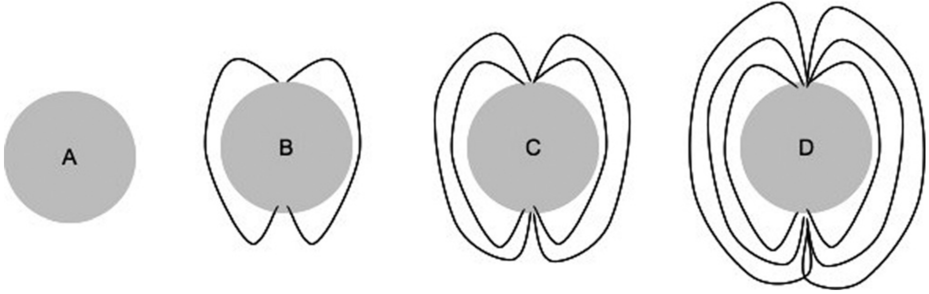
23) Each of the following planets has the same mass M, but each is made of a different material. A darker planet indicates a denser material. Each rocket is at a distance R from the center of each planet. Which rocket experiences a *stronger force* of gravity?

A. A is the strongest because of its greatest density.
B. D is the strongest because of its least density.
C. A is the strongest because it's furthest from the planet.
D. D is the strongest because it's closest to the planet.
E. Gravity is the same for all.



24) Why don't the planets fall into the Sun?

A. Gravity
B. Orbiting is a special case of falling.
C. An orbiting body does not fall.
D. Gravity from other bodies in the solar system is pulling them out.
E. Sun's gravity isn't strong enough to pull them in, only to hold them in place.
F. The planets have their own gravity that counteracts the Sun's.



25) Rank the strength of gravity (greatest to least) on each of the following planets (if any), where more black loops represent stronger magnetic fields. All planets have the same mass and radius.

A. D, C, B, A because stronger magnetic fields make stronger gravity.
B. A, B, C, D because the magnetic fields cancel out a planet's gravity.
C. All have equal ranking because gravity and magnetism have no relationship.
D. There is not enough information to answer the question.

26) Besides gravity, are there any other forces that hold us to the Earth's surface?

A. Atmospheric forces.
B. Rotational dynamics.
C. Friction.
D. Magnetic Forces.
E. More than one of the above factors.
F. Nothing else.

27) How would you compare the strength of gravity at the surface of Pluto with the strength of gravity at the surface of the Earth?
A. Weaker because Pluto has less mass.
B. Weaker because Pluto is further from Sun.
C. Weaker because Pluto has less atmosphere.
D. They are the same.
E. Greater

28) The following planets are viewed from above, with more arrows representing a faster rotation. All planets have the same mass and radius. Rank, *from greatest to least*, the strength of gravity on each planet.
A. C=D, B, A because faster rotation creates more gravity.
B. A, B, C=D because less rotation creates more gravity.
C. All have the same gravity.
D. There is not enough information to answer the question.

For this research project we are requesting demographic information. Due to the make-up of Idaho's population, the combined answers to these questions may make an individual person identifiable. We will make every effort to protect a participants' confidentiality. However, if you are uncomfortable answering any of these questions, you may leave them blank.

29) What is your age?
A) under 18
B) 18-22
C) 23-28
D) 29-33
E) 34-39
F) 40 and over

30) What is your sex?
A) Female
B) Male

31) Do you consider yourself to be a non-native English speaker?
A) No
B) Yes

32) What is your class standing?
A) Freshman
B) Sophomore
C) Junior
D) Senior
E) Graduate

33) What is your major?
A) Science (Physics, Chemistry, Biology, Geology) or Mathematics
B) Engineering
C) Art (Music, Art, Theater, etc.)
D) Humanities (Modern Languages, English, etc.)
E) Social Sciences (Psychology, Sociology, etc.)
F) Business or Economics
G) Education
H) Other

34) What is your physics background?
A) I have never taken physics before
B) I took physics in high school
C) I took another physics course at a university
D) I have taken multiple physics classes

35) How long since your last math class?
A) I am currently enrolled in a math class
B) One semester
C) over 1 year
D) 2 – 5 years
E) over 5 years

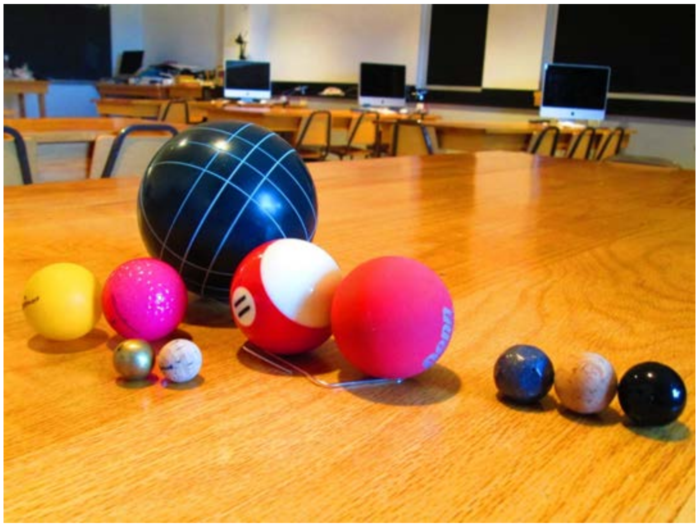
36) What is your highest level of college math completed?
A) I have not taken college math
B) Intermediate Algebra (Math 025)
C) College Algebra (Math 108)
D) Pre-calculus or Trigonometry (Math 143/144)
E) Calculus (Math 160 or 170) or Higher

37) Do you give consent for your answers to be included in the Gravity & Einstein Research study?
A) Yes
B) No

APPENDIX C

Masses Used

Masses were chosen such that it would allow for same mass, but different radius, and same radius, different mass. See list below of the masses used in the lab. They are listed in descending radius length.



Picture 3. Masses used in Lab

1. Bocce Ball: This was chosen because it is hard to find an inexpensive large mass with a reasonably sized radius.
2. Cue Ball
3. Racquet Ball
4. Ping Pong Ball
5. Golf Ball: This has a similar mass to the wooden ball and small brass ball.
6. Lead Ball: Similar in mass to the cue ball.
7. Wooden Ball
8. Black Glass Ball
9. Small Brass Ball
10. Small Cork Ball

APPENDIX D

Tables of Demographic Results

Question/Response		**Pre- Test**			**Post- Test**	
Non-native Engl.	***N***	**M**	**SD**	***n***	**M**	**SD**
No	81	52.3	17.7	66	70.2	17.4
Yes	13	50.5	22.6	19	60.9	20.3
p		**0.66**			**0.06**	
Cohen's *d*		**0.09**			**0.49**	
Effect size correlation *r*		**0.04**			**0.24**	
Class Standing	***N***	**M**	**SD**	***n***	**M**	**SD**
Freshman	33	48	14.2	28	66.4	18.5
Sophmore	29	57.9	21.1	30	68.9	16.6
Junior	20	49.3	16.7	20	69.1	17.9
Senior	8	51.6	19.7	7	60.8	27.1
Graduate	2	49.1	35.4	1^^	96.4	-
p		**0.28**			**0.45**	
Effect size η^2		**0.06**			**0.04**	
Major	***N***	**M**	**SD**	***n***	**M**	**SD**
Sci. or Math.	19	54.9	17.0	19	74.1	15.7
Engin.	9	53.4	20.7	8	66.5	17.6
Art	9	57.8	18.6	7	58.3	10.8
Humanities	5	48.0	14.7	4	66.5	15.9
Soc. Sci.	9	51.2	23.6	7	66.8	23.8
Bus. or Econ.	19	46.1	15.7	20	65.5	20.1
Educ.	6	43.9	9.3	5	60.4	25.6
Other	18	50.5	20.5	16	71.1	18.8
p		**0.70**			**0.58**	
Effect size η^2		**0.05**			**0.07**	

^^ Not included in ANOVA test

Question/Response	**Pre- Test**			**Post- Test**		
Physics Backgrd.	***N***	**M**	**SD**	***n***	**M**	**SD**
None	54	49.0	17.3	49	67.1	18.9
High School	32	53.6	18.8	28	68.5	19.2
Other Univ.	7	58.4	22.5	9	69.7	14.7
Multiple	0	-	-	0	-	-
p		**0.30**			**0.90**	
Effect size η^2		**0.03**			**0.002**	
Time since Math	***N***	**M**	**SD**	***n***	**M**	**SD**
Currently Enrolled	38	51.5	18.4	41	66.6	17.3
One semester	19	45.7	18.0	20	67.9	22.1
over 1 year	17	55.7	15.5	15	71.2	16.5
2-5 years	13	53.8	20.9	8	66.3	19.5
over 5 years	2	43.8	29.0	1^^	95.5	-
p		**0.65**			**0.86**	
Effect size η^2		**0.04**			**0.009**	
Highest Math	***n***	**M**	**SD**	***n***	**M**	**SD**
No college math	9	58.6	18.4	7	63.0	17.7
Interm. Alg.	6	60.6	22.8	4	74.1	12.9
Coll. Alg.	18	50.1	18.2	17	62.5	18.1
Pre-calc. or Trig.	32	51.8	19.0	31	72.6	18.5
Calc. or Higher	23	47.3	15.3	25	66.6	19.6
p		**0.53**			**0.40**	
Effect size η^2		**0.04**			**0.06**	

^^ Not included in ANOVA test

APPENDIX E

IRB Exemption Approval

BOISE STATE UNIVERSITY

DIVISION OF RESEARCH & ECONOMIC DEVELOPMENT
Office of Research Compliance
Institutional Review Board
humansubjects@boisestate.edu | (phone) 208.426.5401 | MS 1138

DATE: **December 18, 2013**

TO: Tiffany Watkins (PI)
Laurie Cavey (co-PI)

FROM: Office of Research Compliance
Institutional Review Board (IRB)

SUBJECT: IRB Notification of Exemption
Project Title: *Gravity & Einstein: Assessing the Rubber Sheet Analogy in Undergraduate Conceptual*

The Boise State University ORC has reviewed your protocol application and has determined that your research is exempt from further IRB review and supervision under 45 CFR 46.101(b).

Review Type: Exempt, Category #1	**Date of Approval: December 18, 2013**
Exemption Approval Number: 024-SB13-134	

This exemption covers any research and data collected under your protocol as of the date of approval indicated above, unless terminated in writing by the principal investigator or the Boise State University IRB. All amendments or changes (including personnel changes) to your approved protocol **must** be brought to the attention of the Office of Research for review and approval before they occur, as these modifications may change your exempt status. Complete and submit a MODIFICATION FORM indicating any changes to your project.

Annual renewals are not required for exempt protocols. When the research project is completed, please notify our office by submitting a FINAL REPORT FORM. The exempt status expires when the research project is completed (closed) or when the review category changes as described above.

All relevant forms are available online. If you have any questions or concerns, please contact the Office of Research Compliance, 208-426-5401 or humansubjects@boisestate.edu.

Thank you and good luck with your research.

Office of Research Compliance